DOCUMENT RESUME

ED 468 192 FL 026 841

AUTHOR	Perron, Veronique, Ed.
TITLE	Actes des Journees de linguistique (Proceedings of the Linguistics Conference) (14th, Quebec, Canada, March 23-24, 2000).
INSTITUTION	Laval Univ., Quebec (Quebec). International Center for Research on Language Planning.
REPORT NO	CIRAL-Pub-B-221
ISBN	ISBN-2-89219-278-1
ISSN	ISSN-1196-121X
PUB DATE	2001-00-00
NOTE	162p.; Produced with the collaboration of Marie-Josee Goulet, Eric Kavanagh, and Caroline Laflamme.
PUB TYPE	Collected Works - Proceedings (021)
LANGUAGE	French
EDRS PRICE	EDRS Price MF01/PC07 Plus Postage.
DESCRIPTORS	Adolescents; Bilingual Education; Canada Natives; English (Second Language); Foreign Countries; French; Grammar; Internet; *Language Research; Language Styles; Language Usage; Linguistic Borrowing; News Media; Phonetics; Phonology; Portuguese; Spanish
IDENTIFIERS	Language Contact; Quebec

ABSTRACT

Papers on language research in this volume include the following: Phonetique ou phonologie: le chva dans les emprunts anglais en portugais (Phonetics or Phonology: The Schwa in English Loan Words to Portuguese) (Pascale Belanger); NewsExpert: exemple de l'application de la linguistique informatique a Internet (NewsExpert: An Example of Its Application to Linguistic Information on the Internet) (Joel Bourgeoys); Proposition d'un classement des genres litteraries: une application de la theorie de l'enonciation (A Proposition of a Classification of Literary Genres: An Application of the Theory of Enunciation) (Sophie Piron); Impact du contact des langues en milieu scolaire a l'ile Maurice (The Impact of Language Contacts in the Academic Setting on the Island of Maurice) (Karina Chavry, with the collaboration of Denise Deshaies); Ceci n'est pas une pipe (This Is Not a Pipe) (Francois Parent); Valeur lexicale et valeur grammaticale (Lexical and Grammatical Value) (Francois Parent); Vues d'aujourd'hui sur le patrimoine (Today's View on Heritage) (Natacha Auclair, Pascale Belinger, Stephanie Campeau, Genevieve Fradette, Michel Jr Laflamme, Marie-France LeBlanc, Patricia Pelletier, Jean-Francois Smith and Marie-Helene Tremblay); La presse quebecoise vue a travers le traitement automatique de ses particularites lexicales (The Press of Quebec and Its Automatic Usages of Its Particular Lexical Items) (Anne Dister); Les expressions partitives en francais et en espagnol (Partitive Expressions in French and Spanish) (Claudine Leullier); education bilingue: consequences sur le comportement linguistiquea l'adolescense (Bilingual Education: How It Effects the Linguistic Use of Adolescents) (Celine Rostiaux); Est-ce que les comparatives et les directionelles sont des series verbales? (Are Comparatives and Directives Verbal Series?) (Renee Lambert); and Langue et identite: perspectives des Inuits (Language and Identity: Perspectives from the Inuits) (Shelley Tulloch). Each paper contains references. (AA)

Reproductions supplied by EDRS are the best that can be made
from the original document.

CENTRE INTERNATIONAL DE RECHERCHE EN AMÉNAGEMENT LINGUISTIQUE

INTERNATIONAL CENTER FOR RESEARCH ON LANGUAGE PLANNING

ED 468 192

Actes des XIVes Journées de linguistique

tenues les 23 et 24 mars 2000 à l'Université Laval

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Publication B-221

Sous la direction de

Véronique Perron

avec la collaboration de Marie-Josée Goulet, d'Éric Kavanagh et de Caroline Laflamme

FL026841

FACULTÉ DES LETTRES

UNIVERSITÉ LAVAL

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

D. Deshaies

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2001

Actes des XIV^{es} Journées de linguistique

tenues les 23 et 24 mars 2000 à l'Université Laval

Sous la direction de
Véronique Perron
avec la collaboration de Marie-Josée Goulet,
d'Éric Kavanagh et de Caroline Laflamme

B-221

2001
CENTRE INTERNATIONAL DE RECHERCHE EN AMÉNAGEMENT LINGUISTIQUE
INTERNATIONAL CENTER FOR RESEARCH ON LANGUAGE PLANNING
QUÉBEC

3

Données de catalogage avant publication (Canada)

Journées de linguistique

Actes des Journées de linguistique

(Publication B)

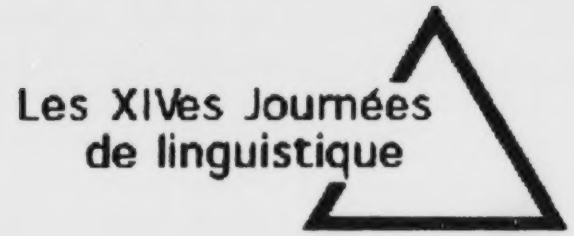
ISSN 1196-121X
ISBN 2-89219-278-1

1. Linguistique – Congrès. 2. Phonétique – Congrès. 3. Français (Langue) – Congrès. 4. Langues en contact – Congrès. I. Association des étudiants diplômés inscrits en langues et linguistique. II. Centre International de recherche en aménagement linguistique. III. Université Laval. Centre international de recherches sur le bilinguisme. IV. Titre. V. Collection: Publication B (Centre international de recherche en aménagement linguistique).

P115.U532 410'.5 C93-031310-0

© **CENTRE INTERNATIONAL DE RECHERCHE EN AMÉNAGEMENT LINGUISTIQUE**
Tous droits réservés. Imprimé au Canada.
Dépôt légal (Québec) – 2e trimestre 2001
ISBN: 2-89219-278-1

Ces actes réunissent les articles des communications présentées lors de la quatorzième édition du colloque étudiant *Les Journées de linguistique* (JDL). Organisé annuellement par l'Association des étudiants diplômés inscrits en langues et linguistique (AÉDILL) de l'Université Laval, les JDL comptent aujourd'hui parmi les plus anciens et les plus importants colloques francophones organisés par et pour des étudiants des cycles supérieurs. La diversité des domaines représentés fait de cette rencontre annuelle, et des actes qui lui succèdent, le point de convergence des recherches les plus actuelles en linguistique.



Actes du colloque sous la direction de Véronique Perron

avec la collaboration d'Annie Bergeron, de Marie-Josée Goulet, d'Éric Kavanagh et de Caroline Laflamme.

Pour une quatorzième fois...

L'organisation d'un colloque, quelle qu'en soit l'ampleur, requiert toujours un investissement considérable en temps, en énergie et en argent. C'est notamment grâce à la participation active de ses membres que l'AÉDILL, l'Association des étudiants diplômés inscrits en langues et linguistique, a pu, encore une fois, réussir le défi d'organiser un événement scientifique de haute qualité et à portée internationale.

Je me joins à mes collègues du Comité organisateur, Éric Kavanagh et Annie Bergeron, pour remercier chaleureusement tous ceux et celles qui ont œuvré à l'organisation des *XIV^es^ Journées de linguistique* (JDL), que ce soit à l'occasion du colloque même ou lors de la préparation des actes qui ont suivi. Nous remercions donc Marie-Josée Goulet, Caroline Laflamme et tous les autres. Nous tenons aussi à remercier tous les participants, notamment les collègues venus de l'extérieur de la région, ainsi que tous les professeurs et chercheurs qui ont assisté à l'événement.

Notre colloque n'aurait pu connaître un tel succès sans l'appui essentiel de ses nombreux commanditaires. Nous remercions tout d'abord l'Association des étudiants de Laval inscrits aux études supérieures et l'Office de la langue française pour leur importante contribution financière. Le Département de langues, linguistique et traduction, le ministère de la Culture et des Communications du Québec et la Faculté des lettres ont été, encore cette année, des alliés de première importance. Nous remercions également la Ville de Sainte-Foy, nouveau commanditaire.

Nous ne saurions assez insister sur le rôle déterminant que joue, année après année, le Centre international de recherche en aménagement linguistique (CIRAL), qui agit, entre autres choses, à titre d'éditeur des actes du colloque. Nous tenons à remercier son directeur, Conrad Ouellon, et aussi Claude Fortin et Dorothy Roy.

Les actes qui suivent sont donc le fruit de plus d'une année d'effort et de travail. Nous vous les présentons avec un double plaisir : celui du travail accompli, d'abord, et celui d'offrir à la linguistique et à nous, linguistes en devenir, une tribune de qualité.

Comme vous le constaterez, les articles colligés ici brossent le portrait d'une linguistique aux mille visages, aux mille accents. Qu'elle porte sur le sens, la forme, le son, le mot ou le texte, qu'elle décrive le français, d'ici ou d'ailleurs, l'inuktitut, le fon, l'espagnol ou le portugais, cette linguistique se présente presque toujours comme l'amalgame fort complexe d'un des objets du discours et d'un cadre théorique. Par ailleurs, forte de ses nouvelles alliances, elle flirte désormais de plus en plus avec les sciences voisines, comme la littérature, la géographie, la sociologie, la psychologie, l'anthropologie, la pédagogie, les sciences de la communication, l'informatique et bien d'autres encore.

Pour une quatorzième fois, merci à tous ces étudiants de linguistique aux cycles supérieurs, et, exceptionnellement, aux quelques étudiants au premier cycle, qui ont participé à l'événement et qui, une fois de plus, assurent le succès et la continuité des JDL.

Véronique Perron
Université Laval, Québec
Le 19 janvier 2001

Table des matières

Avant-propos v

Phonétique ou phonologie : le chva dans les emprunts anglais en portugais 1
Pascale Bélanger
Université Laval

NewsExpert : exemple de l'application de la linguistique informatique à Internet 17
Joël Bourgeoys
Université Laval

Proposition d'un classement des genres littéraires : une application de la théorie de l'énonciation 27
Sophie Piron
Université du Québec à Montréal

Impact du contact des langues en milieu scolaire à l'île Maurice 41
Karina Chavry (avec la collaboration de Denise Deshaies)
Université Laval

Ceci n'est pas une pipe 53
François Parent
Université Laval

Valeur lexicale et valeur grammaticale 65
François Parent
Université Laval

Vues d'aujourd'hui sur le patrimoine 77
Natacha Auclair, Pascale Bélanger, Stéphanie Campeau, Geneviève Fradette, Michel Jr Laflamme, Marie-France LeBlanc, Patricia Pelletier, Jean-François Smith et Marie-Hélène Tremblay
Université Laval

La presse québécoise vue à travers le traitement automatique de ses particularités lexicales 93
Anne Dister
Université de Liège – Université Paris 7

Les expressions partitives en français et en espagnol 107
Claudine Leullier
Université Laval

Éducation bilingue : conséquences sur le comportement linguistique à l'adolescence 119
Céline Rostiaux
Université Libre de Berlin

Est-ce que les comparatives et les directionnelles sont des séries verbales ? 133
Renée Lambert
Université du Québec à Montréal

Langue et identité : perspectives des Inuits 145
Shelley Tulloch
Université Laval

Phonétique ou phonologie : le chva dans les emprunts anglais en portugais

Pascale Bélanger
Université Laval

Plusieurs auteurs soutiennent que l'adaptation des emprunts est phonétique plutôt que phonologique (voir entre autres Silverman, 1992). D'autres tiennent un discours totalement opposé, comme Hyman (1970), Paradis et LaCharité (1997), Paradis, LaCharité et Brault (1999) et Paradis et Prunet (2000), et proposent que l'adaptation des emprunts est phonologique. Ces auteurs apportent de nombreux arguments en faveur de l'adaptation phonologique des emprunts. Un de ces arguments provient du [ə] dans les emprunts anglais en français québécois, lequel est toujours adapté en fonction de sa nature phonologique. En effet, le chva phonétique est adapté selon la voyelle anglaise sous-jacente, et non en fonction de sa nature phonétique. Dans cet article, nous allons montrer que c'est aussi le cas en portugais : le [ə] anglais y est adapté en fonction de la voyelle phonologique sous-jacente lorsqu'il a un statut phonétique et comme une voyelle phonologique entière lorsqu'il a un statut phonologique. Comme nous allons le voir, le comportement de la voyelle chva dans les emprunts anglais en portugais constitue un argument de plus pour l'adaptation phonologique des emprunts.

Emprunts, formes et malformations

La définition de l'emprunt adoptée ici est celle de Paradis (1997), soit un « mot de la langue source qui est incorporé dans le discours de la langue emprunteuse et qui respecte les contraintes les plus périphériques de celle-ci ». *Emprunt* et *forme* ne doivent pas être confondus. Une forme est un emprunt prononcé par le consultant, c'est une réalisation concrète de l'emprunt. Un même emprunt peut

donc avoir plusieurs formes. Les segments de la langue source qui ne sont pas dans l'inventaire phonologique de la langue emprunteuse sont considérés comme mal formés dans cette dernière. Face à un segment mal formé, l'emprunteur a trois choix: l'adaptation du segment selon l'inventaire phonologique de sa langue, l'élision du segment ou encore l'importation de celui-ci.

Dans les corpus d'emprunts du projet de recherche de Carole Paradis et coéquipiers, 84,5% des segments mal formés sont adaptés, 3,5% élidés et 11,8% importés. Les élisions sont rares et ce, en accord avec le principe de préservation, qui stipule que l'information segmentale est préservée maximalement dans les emprunts. Les statistiques générales du projet sur les contraintes en phonologie sont présentées en (1).

(1) Statistiques générales du projet sur les contraintes en phonologie

Corpus	emprunts	formes	malformations analysées	adaptations	élisions	non-adaptations
Emprunts anglais						
français de Québec	948	2412	2080	1659 (79,8%)	79 (3,8%)	342 (16,4%)
français de Montréal	948	2245	1974	1431 (72,5%)	70 (3,5%)	473 (24%)
français de France	300	450	445	364 (81,8%)	18 (4%)	63 (14,2%)
espagnol mexicain	1045	1514	2108	1405 (66,6%)	1 (0,1%)	661 (31,4%)*
Emprunts français						
arabe marocain	1127	2691	3941	3071 (77,9%)	351 (8,9%)	519 (13,2%)
kinyarwanda	756	2130	4208	4120 (97,9%)	62 (1,5%)	26 (0,6%)
lingala	672	1917	3571	3378 (94,6%)	31 (0,9%)	162 (4,5%)
peul	545	1036	814	741 (91%)	57 (7%)	16 (2%)
Total 8 corpus	6341	14395	19 141	16 169 (84,5%)	669 (3,5%)	2262 (11,8%)*

* Il faut ajouter aux pourcentages de cette ligne 41 cas marginaux, soit 1,9% en espagnol mexicain et 0,2% au total des corpus.

ARGUMENTS EN FAVEUR DE L'ADAPTATION PHONOLOGIQUE

Le comportement du /r/ français en arabe marocain

Paradis, LaCharité et Brault (1999) ont apporté de nombreux arguments en faveur de l'adaptation phonologique des emprunts. Un de leurs arguments provient du comportement de la rhotique /r/ dans les emprunts français en arabe marocain.

Béland, Paradis et Bois (1993) ont étudié le comportement de la rhotique française dans une étude de groupe sur les paraphasies phonémiques en français. Ils montrent que la rhotique française se comporte phonologiquement comme une coronale. Phonétiquement, le phonème /r/ connaît de nombreuses variantes en français. Elle peut être réalisée comme une vibrante apico-alvéolaire, une uvulaire roulée, une uvulaire fricative et même comme une laryngale, toutes ces variantes pouvant aussi être dévoisées, comme il est très fréquent en fin de mot. Malgré les différences acoustiques et articulatoires importantes entre les variantes, les locuteurs francophones n'éprouvent aucune difficulté à reconnaître qu'il s'agit toujours de la même unité phonologique.

Les arabophones reconnaissent aussi que les réalisations phonétiques du /r/ sont des allophones d'un même phonème en français. Dans le corpus d'emprunts français en arabe marocain, la consonne /r/ est systématiquement réalisée [r]. Par exemple, l'emprunt français *rouleau*, habituellement prononcé [ʁulo] en français, est adapté en [rulo], comme présenté en (2).

(2)	français		arabe marocain
rouleau	[ʁulo]	→	[rulo]

On pourrait penser que les locuteurs arabes effectuent cette catégorisation parce qu'ils n'entendent pas la réalisation phonétique

uvulaire de [ʁ]. Au contraire, les arabophones sont aptes à entendre les réalisations post-vélaires puisque leur système phonologique, qui comprend la consonne uvulaire /ʁ/, exploite cette région articulatoire. Malgré la réalisation phonétique uvulaire et la présence d'une unité phonologique de même nature dans leur système, les arabophones catégorisent l'uvulaire phonétique française comme une coronale.

Inversement, la fricative uvulaire /ʁ/ arabe est considérée comme mal formée en français lorsqu'elle est contenue dans un emprunt arabe. En effet, elle est systématiquement adaptée en /g/, comme dans le mot *gazelle*, présenté en (3).

(3)	arabe classique		français
gazelle	[ʁazæ:l]	→	[gazɛl]

L'uvulaire /ʁ/, distinctive en arabe, est considérée comme mal formée en français et ce, même si les locuteurs francophones sont aptes à percevoir et à produire la variante phonétique [ʁ], qui possède les mêmes caractéristiques acoustiques et sonores que la consonne phonémique arabe.

L'adaptation de la rhotique française /ɾ/ en arabe ainsi que de la rhotique arabe /ʁ/ en français indique que les variantes phonétiques n'influencent pas l'adaptation phonologique des emprunts.

Le comportement de la glide française /ɥ/ en peul

Un deuxième argument provient du comportement du segment /ɥ/ dans les emprunts français en peul. La glide française /ɥ/ est absente de l'inventaire phonologique peul. Par conséquent, elle est considérée comme mal formée dans cette langue.

Pourtant, [ɥ] est une variante phonétique possible de /w/ en peul. En effet, la glide phonémique peule /w/, lorsque devant les voyelles d'avant [i] et [ɛ], peut être réalisée [ɥ]. Par exemple, le mot peul /wɛccɔ/, qui signifie « poitrine », peut être prononcé [ɥɛccɔ] dans certaines variétés de peul. Par contre, même si la variante phonétique [ɥ] existe en peul, le phonème français /ɥ/ est considéré comme mal formé et est élidé dans les emprunts. Comme on le voit en (4), dans l'emprunt français *tuyau* [tɥijo], même si l'environnement phonétique est favorable à la réalisation phonétique du [ɥ] peule, /ɥ/ est élidé. Ici, l'élision est provoquée par une double contrainte. Celle-ci s'explique par le fait que /ɥ/ est imbriqué dans une suite syllabique mal formée en peul, qui ne permet par les attaques branchantes.

(4)	français		peul
tuyau	[tɥijo]	→	[tijo]

Le comportement de la laryngale /h/ anglaise dans certains dialectes québécois

Un troisième argument provient du comportement des fricatives alvéolaires /ʃ/ et /ʒ/ françaises, qui sont prononcées [h] dans certaines régions du Québec. Par exemple, le mot français *changer* /ʃɑ̃ʒe/ peut être prononcé [h̲ɑ̃h̲e]. Cette variante phonétique n'empêche pas la consonne laryngale phonémique /h/ anglaise d'être considérée comme mal formée et d'être élidée dans les emprunts anglais, au même titre que dans toutes les autres régions du Québec. Par exemple, en (5), on voit que l'emprunt anglais *hamburger* est prononcé [ambəgəɹ] (corpus de Québec). L'élision du /h/ est effectuée selon le principe de la non-disponibilité, qui stipule qu'« un primitif phonologique non disponible dans la langue

emprunteuse ne peut être phonologiquement traité par cette langue[1] ».

(5)	anglais		français québécois
hamburger	[hæmbəɹgəɹ]	→	[ambəɹgœɹ]
			[ambəɹgəɹ]
			[ambəgəɹ]

Ces trois cas montrent bien que l'adaptation des emprunts est phonologique et non phonétique. Le but de cet article est d'apporter un autre argument à l'adaptation phonologique des emprunts, et ce, à partir du comportement de la voyelle anglaise *ə* dans les emprunts anglais en portugais. Comme nous le verrons, la voyelle est toujours prononcée selon la voyelle sous-jacente lorsqu'elle a un statut phonétique et adaptée de façon stable lorsqu'elle a un statut phonologique.

Méthodologie

Les emprunts anglais constituant notre corpus ont été recueillis dans divers ouvrages, dont *Estrangeirismos na Língua Portuguesa* de José Pedro Machado (1994) et le *Dicionário dos anglicismos e germanismos na língua portuguesa* de Radefeld Schmidt (1997). Ils ont été prononcés par trois Portugais originaires de la région de Minho, au nord de Porto, à partir d'un questionnaire écrit. On pourrait objecter que la graphie pourrait influencer les adaptations (étant donné la méthode utilisée), mais des études sur l'influence de la graphie ont démontré que celle-ci n'influence que très faiblement nos corpus[2].

[1] Voir Paradis, LaCharité et Brault (2000).

[2] Au total, l'influence de la graphie n'est que de 3,7%.

Les emprunts prononcés par les consultants ont été enregistrés, puis la prononciation de deux consultants a été transcrite en alphabet phonétique international[3] (API) et informatisée dans une base de données[4]. Chaque forme d'emprunt qui contenait la voyelle mal formée ə a été analysée.

L'inventaire vocalique des langues en contact

L'inventaire vocalique portugais, présenté en (6), comprend sept phonèmes.

(6) Inventaire phonologique des voyelles du portugais européen excluant les diphtongues (adapté de Mateus, 1982)

	-arrière				+arrière			
	-arrondi		+arrondi		-arrondi		+arrondi	
	+tendu	-tendu	+tendu	-tendu	+tendu	-tendu	+tendu	-tendu
+haut	i						u	
-haut - bas	e	ɛ					o	ɔ
+bas					a			

Le système vocalique phonologique de l'anglais utilisé est présenté en (7). Notons que le chva y est classé comme une voyelle d'arrière[5].

[3]Les transcriptions phonétiques effectuées ne sont pas étroites. Entre autres, la version vélarisée de [l], la longueur des voyelles ainsi que le flap n'ont pas été transcrits parce que très prévisibles en anglais.

[4]Le temps écoulé entre la construction du corpus et la présentation de ces recherches ne nous a pas permis de faire les transcriptions phonétiques des formes du 3[e] consultant.

[5]Myriam Naud, lors d'une communication aux XIVes Journées de linguistique, a présenté des arguments en faveur du classement de la voyelle anglaise chva comme une voyelle d'arrière.

(7) Inventaire phonologique des voyelles de l'anglais excluant les diphtongues (adapté de Jensen, 1993)

	-arrière				+arrière		
	-arrondi		+arrondi		-arrondi	+arrondi	
	+tendu	-tendu	+tendu	-tendu	+ / - tendu	+tendu	-tendu
+haut	i	ɪ				u	ʊ
-haut - bas	e	ɛ			ə	o	
+bas	æ						ɑ

5 des 10 voyelles anglaises sont mal formées en portugais, soit /æ/, /ɪ/, /ə/, /ʊ/ et /ɑ/. Les voyelles anglaises mal formées sont ombragées dans le tableau (7).

Le chva phonétique

Le ə n'a pas toujours le même statut en anglais. Lorsqu'il est la réduction d'une voyelle pleine inaccentuée, il a un statut purement phonétique. Ce phénomène de réduction vocalique, qui pose qu'une voyelle pleine contenue dans une syllabe non accentuée est souvent réduite à [ə], est très fréquent en anglais. Quelques exemples de réduction vocalique sont présentés en (8).

(8)	i → ə	*nativity* 'nativité'	[netɪ̲vəti]	vs	*native* 'natif'	[nétə̲v]
	ɑ → ə	*symbolic* 'symbolique'	[sɪmbɑ̲́lɪk]	vs	*symbol*' 'symbole'	[símbə̲l]
	ɛ → ə	*less* 'sans'	[lɛ̲́s]	vs	*helpless* 'sans recours'	[hɛ́lplə̲s]
	æ → ə	*fatality* 'fatalité'	[fetǽ̲ləti]	vs	*fatal* 'fatal'	[fétə̲l]

Dans notre corpus, on dénombre un total de 369 cas où ´ est phonétique. De ce nombre, les adaptations représentent 79,7% des cas (294 / 369), les importations 18,2% (67 / 369), et les élisions 2,1% (8 / 369).

Pour cet article, nous allons nous concentrer sur les 294 cas d'adaptation. L'objectif est de voir si l'emprunteur prononce [ə] en fonction de la voyelle phonologique sous-jacente ou s'il ne tient compte que de la manifestation phonétique du [ə].

Selon l'hypothèse de l'adaptation phonologique, l'emprunteur, que nous considérons comme bilingue[6], a accès au système phonologique de la langue source. Il devrait donc être en mesure de restituer (de façon directe ou indirecte) la voyelle pleine sous-jacente à [ə]. Par conséquent, le patron d'adaptation devrait être variable.

Cette hypothèse se vérifie en portugais, où sept patrons d'adaptation différents sont observés pour le chva phonétique, comme on le voit en (9). Les différentes voyelles se partagent inégalement le total des cas (de 1,1% à 34,4% de ceux-ci).

[6]Sur la base de résultats phonologiques et sociolinguistiques, Paradis et LaCharité (1997, 2000), Prunet (1999) et Paradis et Prunet (2000) posent explicitement que les emprunteurs sont des bilingues, lesquels ont accès à la langue source. Nous adoptons cette position.

(9) Résultats

Patron d'adaptation	*% des cas analysés*
ə → ɛ	34,4 % (96 / 279)
ə → a	30,8 % (86 / 279)
ə → ɔ	17,5 % (49 / 279)
ə → i	13,3 % (37 / 279)
ə → o	2,9 % (8 / 279)
ə → e	1,1 % (3 / 279)
ə → u	3,4 % (cas éliminés)

Les cas dans lesquels [ə] devient [u], soit 3,4% des cas, ont été éliminés. Il existe un processus natif en portugais qui réduit systématiquement /ɔ/ et /o/ inaccentués à [u]. Lorsque [ə] devient [u], il est difficile de savoir s'il s'agit d'une adaptation phonologique du résultat de la règle native en question, ou encore d'une influence analogique. C'est pourquoi ces cas ont été éliminés.

On peut déterminer la voyelle sous-jacente au [ə] inaccentué dans les emprunts à l'aide des dérivés morphologiques de ceux-ci. Quelques exemples de prononciation sont présentés en (10).

(10) Exemples de prononciation de la voyelle [ə]

anglais		*mot apparenté*		*portugais*
a. ə → ɛ *topless* [tɑplə̲s]	'topless'	*less* [lɛ̲́s]	→	[tɔ́plɛ̲s]
b. ə → a *pentagon* [pɛ́ntə̲gɑ̀n]	'pentagone'	*pentagonal* [pɛntǽ̲gənəl]	→	[pɛntá̲gɔn] [pɛntá̲gɔno]
c. ə → ɔ *sex symbol* [sɛ́ksɪ̀mbə̲l]	'sexe symbole'	*symbolic* [sɪmbɔ̲́lɪk]	→	[sɛksímbɔ̲l]
d. ə → i *festival* [fɛ̃stə̲vəl]	'festival'	*festivity* [fɛstɪ̲́vəti]	→	[fɛstɪ̲́val]
f. ə → e *star company* [stɑ́ɹkəmpə̲ni]	'star company'	*companion* [kəmpǽ̲njən]	→	[stɑ́ɹkɔ́mpɛ̲ni]
g. ə → o *astronaut* [ǽstrə̲nɑ̀t]	'astronaute'	*astronavigation* [æstɹɔ̲́nævəgəʃən]	→	[ástro̲nawt]

Par exemple en (10b), [ə] devient [a] dans l'emprunt *pentagon* [pɛ́ntə̲gɑ̀n]. À l'aide du dérivé *pentagonal* [pɛntǽ̲gənəl], on peut déduire que le chva inaccentué est la réduction d'un /æ/ phonologique. Il est aussi adapté comme tel : [pɛnta̲gɔn] ou [pɛnta̲gɔno] selon le consultant, qui restitue indirectement la voyelle sous-jacente sous la forme de /a/. Cette adaptation, basée sur une modification du trait [arrière], est minimale, et cela en accord avec le principe de minimalité.

Autre exemple en (10c), où [ə] devient [ɔ] dans l'emprunt *sex symbol* [sɛ́ksɪ̀mbə̲l]. Selon le dérivé *symbolic* 'symbolique' [sɪmbɔ̲́lɪk], le chva inaccentué de *sex symbol* est la réduction d'un /ɑ/ phonologique. Il est aussi adapté comme tel : l'emprunt devient [sɛksimbɔ̲l] en portugais, la voyelle de base est ainsi indirectement

restituée sous la forme de /ɔ/. Cette adaptation, qui est aussi minimale, est basée sur le trait de hauteur.

Sept patrons différents sont donc observés en portugais pour l'adaptation du chva phonétique (voir 9). Selon nos résultats, les emprunteurs restituent la voyelle sous-jacente au [ə], sur laquelle ils basent leur adaptation.

Le chva phonologique

Le chva peut aussi avoir un statut phonologique, entre autres lorsqu'il est en position accentuée. Dans notre corpus, le chva est phonologique dans 297 cas. Il connaît un taux d'importation très élevé, soit 62,9% (187 / 297 cas), contre 37,1% d'adaptation (110 / 297 cas). Notons que 100% des /ə/ importés sont contenus dans le suffixe [-əɹ], un suffixe morphologique très productif en anglais qui serait, selon nos résultats, relativement bien intégré en portugais. Les segments qui le composent seraient dans la périphérie phonologique portugaise (voir Paradis et Lebel (1997) pour un exposé détaillé sur la périphérie dans les emprunts).

Si l'on considère le statut phonologique de /ə/, on devrait s'attendre à un patron d'adaptation différent de celui de [ə]. Puisque les emprunteurs sont bilingues et qu'ils ont accès à la phonologie de la langue source, ils devraient avoir conscience du statut phonologique de /ə/. Par conséquent, le patron d'adaptation devrait être beaucoup plus stable.

C'est exactement ce que l'on observe en portugais, où un patron d'adaptation domine nettement, comme présenté en (11).

(11) Résultats

Patron d'adaptation	*% des cas analysés*
ə → ɔ	94,1 % (80 / 85 cas)
cas déviants*	5,9 % (5 / 85 cas)

*Les cas déviants sont les suivants:
ə → ɛ *punch* [pə́ntʃ] → [pɛ́ʃ], *punk* [pə́ŋk] → [pɛ́k], *rugby* [ɹə́gbi] → [ɹɛ́gbi]
ə → a *lunch* [lə́ntʃ] → [lá], *tupperware* [tə́pəɹwɛ̀ɹ] → [tapperwer]

Dans 94,1% des cas, les consultants ont adapté /ə/ en /ɔ/, une voyelle d'arrière. Par exemple, le mot anglais *club* [klə́b] est prononcé [klɔ́b] en portugais, comme on le voit en (12).

(12)	*anglais*		*portugais*	
bluff	[blə́f]	→	[blɔ́f]	'bluff'
club	[klə́b]	→	[klɔ́b]	'club'
network	[nɛtwə́ɹk]	→	[nɛtwɔ́ɹk]	'réseau'

La voyelle phonologique /ə/ n'est donc pas catégorisée comme la variante phonétique [ə]. Cela montre que les emprunteurs reconnaissent le statut phonologique de celle-ci puisqu'ils l'adaptent de façon stable dans 94,2% des cas (80/85). Seulement 5,8% des cas (5/85) dévient du patron d'adaptation principal. Les deux adaptations déviantes sont aussi minimales, la première implique une transformation de la valeur du trait de hauteur, la seconde l'ajout d'un nœud labial. Ces stratégies peuvent être soit le fruit d'une influence analogique ou de l'influence d'une autre langue, mais certainement pas de l'influence orthographique. Malgré ces quelques cas déviants, les résultats démontrent clairement que les emprunteurs, qui sont bilingues et qui ont accès à la phonologie de la langue source, connaissent le statut phonologique de la voyelle /ə/.

En conclusion, deux stratégies d'adaptation extrêmement différentes sont observées selon que le chva est phonétique ou phonologique. La première stratégie, basée sur la voyelle sous-jacente, suggère un patron d'adaptation variable, alors que la seconde, qui résulte d'un patron d'adaptation stable, est basée sur la nature phonologique de chva. Ces résultats, selon lesquels les emprunteurs connaissent le matériel phonologique de la langue source, constituent un argument de plus pour l'adaptation phonologique des emprunts.

Bibliographie

BÉLAND, R., C. PARADIS, et M. BOIS (1993), « Constraints and repairs in aphasic speech : A group study », dans C. PARADIS et D. LACHARITÉ (dir.), *Constraint-based theories in multilinear phonology,* numéro spécial, *Revue canadienne de linguistique* 38,2, p. 279-302.

HYMAN, L. (1970), « The role of borrowing in the justification of phonological grammar », *Studies in African Linguistics* 1(1), p. 1-48.

JENSON, J. (1993), *English phonology*, Amsterdam, John Benjamins Publishing Company.

MACHADO, J. P. (1994), *Estrangeirismos na Língua Portuguesa*, Éditorial Notícias, Lisboa, Colecção Linguistica.

MATEUS, M. H. (1982), *Aspectos da Fonologia Portuguesa*, Instituto Nacional de Investigação Científica.

PARADIS, C. (1997), « Préservation phonémique et seuil de tolérance dans l'adaptation d'emprunts », dans J. AUGER et Y. ROSE (dir.), *Explorations du lexique*, Québec, CIRAL, p. 215-228.

PARADIS, C., et D. LACHARITÉ, *Guttural deletion in doanwords*. Manuscrit.

PARADIS, C., et D. LaCharité (1997), « Preservation and Minimality in Loanword Adaptation », *Journal of Linguistics*, 33,2, p. 379-430.

PARADIS, C., D. LACHARITÉ, et F. BRAULT (1999), « L'élision extraordinaire des gutturales pharyngales et laryngales dans les emprunts et le principe de la non-disponibilité », *Revue canadienne de linguistique* 44,2. Manuscrit.

PARADIS, C., et É. LEBEL (1997), « Centre et périférie : les segments anglais en français québécois », dans J. AUGER et Y. ROSE (dir.), *Explorations du lexique*, Québec, CIRAL, p. 229-242.

PARADIS, C., et J.-F. PRUNET (2000), « Nasals voyels as two segments : Evidence from borrowings ». À paraître dans *Language* 76, 2. Manuscrit.

SCHMIDT, R. (1997), *Dicionário dos anglicismos e germanismos na língua portuguesa*, Frankfurt am Main, TFM.

SILVERMAN, D. (1992), « Multiple scansions in loanword phonology : Evidence from Cantonese », *Phonology* 9, p. 289-328.

NewsExpert : exemple de l'application de la linguistique informatique à Internet

Joël Bourgeoys
Université Laval

Introduction

Nous traitons dans cet article d'une application de la linguistique informatique à Internet. Il est question d'un logiciel conçu par Frédérick Brault, Marie-Pierre Hétu et moi-même, Joël Bourgeoys. Ce logiciel fut construit dans le cadre d'un séminaire de linguistique informatique donné par Jacques Ladouceur, professeur à l'Université Laval. L'objectif de départ était de créer un logiciel qui comporterait un composant linguistique. Ce logiciel devait pouvoir être utilisé par quiconque et, à la limite, être commercialisable. Nous vous présentons donc certaines étapes de la construction d'un logiciel, les problèmes rencontrés, ainsi que les solutions à ces problèmes.

Les groupes de discussion

Le fonctionnement des groupes de discussion, aussi appelés forums, est semblable à celui du courrier électronique. La différence principale est que les messages envoyés à un groupe de discussion peuvent être lus par un nombre illimité de lecteurs. Ces groupes comportent donc souvent plusieurs messages, et ce, à un tel point qu'il devient fastidieux de lire les messages pour trouver de l'information qui pourrait nous intéresser.

Les groupes de discussion sont très variés. Il existe, sur Internet, plus de 40 000 groupes de discussion qui traitent de presque tous les sujets imaginables, des problèmes cardio-vasculaires à la culture

québécoise en passant par les sports, les voitures et plusieurs émissions de télévision.

Alors que certains groupes reçoivent un ou deux messages par semaine, d'autres peuvent en recevoir des milliers par jour. C'est alors que la gestion des messages devient ardue. Lire une centaine de messages est une tâche qui prend beaucoup de temps.

Les logiciels de lecture des groupes de discussion

Plusieurs logiciels existent déjà pour la lecture des groupes de discussion. Ils fonctionnent presque tous de la même manière. Il y a habituellement trois fenêtres : une affichant les groupes de discussion auxquels l'utilisateur est abonné, une autre affichant les en-têtes et, finalement, une dernière comportant le corps du message. Habituellement, les différents logiciels téléchargent les en-têtes des messages et les insèrent dans un arbre dont le tronc comporte les messages principaux et les branches, les réponses à ces messages. Ainsi, l'utilisateur peut avoir une vue d'ensemble des messages existant dans ce groupe de discussion.

La plupart des logiciels offrent aussi la possibilité de faire une recherche par mot-clef à l'intérieur des messages. La recherche par mot-clef implique une question bien précise : Qu'est-ce qu'on recherche? Nous croyons par contre que la recherche d'information précise n'est pas le but premier des utilisateurs de groupes de discussion. Les groupes de discussion sont, en quelque sorte, comme les quotidiens : on n'achète pas le quotidien *Le Soleil* pour voir si la situation linguistique en Afrique y est discutée, mais plutôt pour en apprendre sur des sujets divers.

Problèmes liés à la rubrique « sujet »

Nous avons vu que les messages étaient le plus souvent classés selon la rubrique « sujet ». Nous croyons cependant que cette méthode de classification pourrait être améliorée, et ce, pour deux raisons. La première étant que l'information contenue dans la rubrique sujet est souvent non représentative du contenu du message. Par exemple, dans un groupe de discussion traitant des langages de programmation, un message pourrait avoir comme en-tête « à l'aide » ou encore « besoin d'aide avec Prolog ». Dans ces deux cas, il est impossible pour quiconque de savoir exactement sur quoi porte le message en question. À la limite, nous savons que le deuxième porte sur le langage Prolog, mais le message doit être lu pour savoir de quel problème il s'agit. Le sujet réel du message n'est donc pas nécessairement explicité dans l'en-tête. La deuxième raison, qui est rattachée à la première, est que le message peut contenir d'autres informations pertinentes en plus de celle fournie dans l'en-tête sujet. Par exemple, un message dans un groupe discutant des voitures Volkswagen pourrait avoir comme en-tête « performances de la Golf » mais discuter aussi de la Jetta et de la Passat.

Il devient alors évident que l'organisation des messages par l'en-tête pourrait être améliorée. L'utilisateur doit soit lire tous les messages, soit les passer rapidement en prenant la chance de laisser de côté de l'information qui pourrait l'intéresser. Comment alors améliorer l'organisation des messages? Puisque les messages renferment des données textuelles, il est possible d'en faire un traitement linguistique. Nous recherchons, pour remplacer la rubrique sujet, les thèmes des messages, c'est-à-dire les termes utilisés qui sont pertinents aux messages. Nous avons donc mis au point un logiciel qui télécharge les en-têtes et les messages dans les groupes de discussion, les analyse, et retire les mots importants de ces messages.

Généralités de NewsExpert

NewsExpert est différent des autres lecteurs de messages. Il ne fait pas qu'afficher les messages et permettre une recherche par mot-clef. Il fait aussi une classification des messages basée sur l'information que comporte le texte.

L'interface du logiciel est divisée en quatre grandes fenêtres. La première affiche les groupes de discussion auxquels l'utilisateur est abonné, la deuxième affiche la rubrique « en-tête » des différents messages, la troisième donne les thèmes principaux contenus dans les messages et finalement, la quatrième affiche le texte même du message.

NewsExpert permet d'abord les opérations de base comme quitter le logiciel, imprimer et se connecter à un serveur, par exemple. La connexion au serveur contient les choix offerts par la plupart des logiciels : le nom du serveur, le port, le nom d'usager et le mot de passe. Il est aussi possible de régler la langue de l'interface du logiciel. Jusqu'à maintenant, les langues disponibles sont le français, l'anglais et l'espagnol.

Après s'être connecté à un serveur, l'utilisateur a l'option d'aller consulter les différents groupes de discussion disponibles sur ce serveur. Il doit, pour ce faire, télécharger le nom des groupes auxquels il peut ensuite s'abonner. Après s'être abonné à un ou plusieurs groupes de discussion, le logiciel est prêt pour l'analyse des messages. Le logiciel commence par télécharger les en-têtes des messages. Ensuite, il télécharge le corps des messages et les insère un à la suite de l'autre pour former le corpus. Le corpus est découpé en mots qui sont placés dans des bases de données. Le logiciel analyse ensuite ces données pour retenir les mots les plus importants. L'analyse s'effectue en arrière-plan, et ce, sans que l'utilisateur ne s'en rende compte.

Calcul fréquence-distribution

L'importance d'un mot est basée sur sa fréquence et sa distribution. Grâce à la loi du moindre effort, il est possible de classifier l'importance des mots par leur fréquence (Salton, 1983 : 60). En effet, l'auteur d'un texte utilise souvent les mêmes mots pour parler des mêmes choses. Ainsi, nous pourrions supposer que plus un mot est utilisé souvent, plus il est un descripteur important. Cependant, nous savons aussi que les mots-outils ont une fréquence importante dans un texte. En observant les mots à l'intérieur d'un texte, on se rend compte que les mots les plus fréquents sont souvent les mots-outils, on retrouve ensuite des mots plus importants et finalement, les mots qui sont moins significatifs se retrouvent à la fin de la liste. C'est donc dire que les mots que nous cherchons se trouvent au milieu. Cependant, nous retrouvons des mots non significatifs à toutes les fréquences. C'est pourquoi il est aussi utile d'utiliser la distribution pour calculer l'importance d'un mot.

La distribution d'un mot est définie par le nombre de messages dans lesquels ce mot apparaît. Si un mot apparaît dans un seul message, ce mot peut être un bon descripteur, mais uniquement pour ce message. Il n'est donc pas intéressant pour nos besoins. Par contre, un mot se trouvant dans tous les messages n'est pas plus intéressant parce qu'il ne décrit pas une partie des messages mais plutôt tous les messages. Par exemple, le mot *le* se retrouverait probablement dans tous les messages. Ici encore, il y a un juste milieu à trouver.

En rassemblant la fréquence et la distribution, nous pouvons donc conclure que le mot idéal est celui dont la fréquence est élevée, mais la distribution moyenne. Ainsi, les mots qui ont une fréquence et une distribution trop élevée ou trop basse sont retirés de la liste. Un calcul est ensuite effectué avec la fréquence et la distribution afin de déterminer l'importance des mots.

Le thème

Après l'analyse, le logiciel affiche les thèmes retenus dans un arbre. Ces termes sont placés en ordre d'importance. L'utilisateur peut donc faire défiler la banque de thèmes et choisir celui ou ceux qui l'intéressent. En cliquant sur un thème, le logiciel affiche, dans une autre fenêtre, les en-têtes des messages qui contiennent ce thème. Par exemple, dans un groupe de discussion traitant de la biologie, l'utilisateur choisit le thème *conference*, trois messages sont alors retenus, qui ont en commun le thème de colloques en biologie. Ensuite, en cliquant sur une en-tête de message, l'utilisateur voit le texte du message correspondant s'afficher. De cette façon, il peut lire tous les messages comportant ce thème.

Les sous-thèmes

Le logiciel va encore plus loin. Il est fort possible, dans un groupe de discussion volumineux, que le thème intéressant se trouve dans une vingtaine de messages, voire davantage. L'utilisateur qui n'est pas intéressé par tous ces messages peut double-cliquer sur le thème. Ceci a pour effet d'extraire ce que nous avons appelé les sous-thèmes. Les messages comportant le premier thème choisi sont alors analysés afin de faire ressortir les mots pertinents à ces 20 messages. Ceci a pour avantage de permettre à l'utilisateur de « naviguer », en quelque sorte, à l'intérieur des messages.

Si nous revenons à l'exemple précédent, les messages qui avaient pour thème *conference* avaient été retenus. Si l'utilisateur voulait maintenant observer les « sous-thèmes » appartenant à ces messages, il n'aurait qu'à double-cliquer sur le thème *conference* pour avoir accès à la banque de sous-thèmes. L'utilisateur, alors intéressé par le sous-thème *computer*, n'aurait qu'à cliquer sur le sous-thème pour voir apparaître deux messages. Les deux messages ont alors comme thème principal *conference* et comme sous-thème *computer*. L'utilisateur peut alors lire les

messages qui ont comme sujet les conférences ayant un lien quelconque avec l'informatique et la biologie : « Artificial Neural Networks In Medicine And Biology » et « The 2000 International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences ».

Il devient alors évident que l'option des sous-thèmes est très pratique lorsqu'il ne reste qu'une cinquantaine de messages à lire, après avoir sélectionné un thème.

Problèmes rencontrés

Le logiciel s'inscrit dans le domaine de la recherche automatisée d'information. Ce domaine a pour but d'aider l'utilisateur à trouver l'information qu'il désire parmi un grand nombre de documents. Le logiciel que nous avons créé est basé sur des statistiques de corpus. Soumis à de gros corpus, l'application répond à nos attentes, et permet de retenir les termes importants. Cependant, nous obtenons beaucoup de bruit parmi les thèmes, c'est-à-dire que plusieurs des mots retenus ne sont pas significatifs.

Nous avons donc fait quelques tests avec des anti-dictionnaires afin d'éliminer certains mots-outils qui se retrouvaient souvent dans nos corpus. Un anti-dictionnaire a pour but d'éliminer certains termes (Salton, 1983 : 71). Par exemple, si un anti-dictionnaire comporte les mots *le*, *la*, *un* et *une*, ceux-ci ne seront pas traités lors de l'analyse. Nous en sommes arrivés à la conclusion que l'utilisation d'anti-dictionnaires améliore les résultats. Par contre, l'utilisation d'un anti-dictionnaire amène aussi une certaine difficulté. L'analyse statistique d'un corpus est indépendante de la langue. L'analyseur peut traiter un texte rédigé dans n'importe quelle langue en autant que les caractères soient latins et que les mots soient séparés par des espaces pour le découpage des mots. Cependant, en ajoutant un anti-dictionnaire, l'analyseur devient dépendant d'une langue en particulier. Il serait peut-être avantageux

d'observer si les calculs statistiques ne pourraient pas être améliorés avant d'ajouter des anti-dictionnaires.

Lors d'essais du logiciel, nous nous sommes rendus compte que certains thèmes apparaissaient plus d'une fois. Par exemple, dans un groupe discutant de biochimie, les thèmes *computer*, *computers* et *computing* étaient présentés dans la fenêtre des thèmes. Ces trois mots faisaient cependant partie du même thème. C'est pourquoi nous avons décidé d'utiliser un lemmatiseur. Un lemmatiseur ramène les variantes à une forme commune. Le verbe *mangerais* serait donc réduit au mot *manger*, tout comme *mangeons* et *mangerons*. Comme pour l'anti-dictionnaire, nous ne voulions pas ajouter un module qui rendrait le logiciel dépendant d'une certaine langue. Nous avons donc utilisé un calcul de similitude. Il s'agit tout d'abord de faire la compilation des doublons dans les deux mots. Par exemple, le mot *livre* contient les doublons suivants: *li*, *iv*, *vr* et *re*. Le mot *livres*, lui, contient *li*, *iv*, *vr*, *re*, *es*. Nous devons ensuite calculer le pourcentage des doublons d'un mot qui se retrouvent dans l'autre mot. Ici, les doublons du mot *livre* se retrouvent à 100% dans le mot *livres* alors que 4 doublons sur 5 (80%) du mot *livres* se retrouvent dans le mot *livre*. En faisant la moyenne des deux (100% + 80% / 2), nous atteignons un taux de similitude de 90%. Une des particularités de ce calcul est de traiter aussi les erreurs. Le système fonctionne de la même manière. Si nous prenons l'exemple du mot *traduction* et de l'erreur *traducsion*, sept doublons sur neuf sont identiques dans les deux cas. Ainsi, nous avons un taux de similitude de 77%. Le taux de similitude satisfaisant n'est pas coulé dans le béton. Il s'agit de faire des tests et d'observer quel taux de similitude est adéquat pour le logiciel. Nous avons fixé le seuil de similitude à 60%.

D'autres fonctionnalités amélioreraient aussi NewsExpert. Premièrement, le logiciel devrait reconnaître les termes complexes. Par exemple, le nombre 1991 n'a pas une grande signification en lui-même sauf lorsqu'il s'agit d'une année. Cependant, les termes *Honda 1991* désignent une voiture particulière. Nous croyons donc

qu'avec les termes complexes, les thèmes retenus seraient encore plus pertinents.

Conclusion

Nous avons créé un logiciel de gestion des messages de groupes de discussion qui fonctionne très bien. Quelques fonctionnalités devraient être ajoutées mais nous croyons que l'idée de départ est bonne et que notre logiciel pourrait servir les gens qui consultent les groupes de discussion volumineux.

NewsExpert n'est qu'un exemple de l'apport de la linguistique à la gestion de l'information sur Internet. Les mêmes principes pourraient, par exemple, aussi bien s'appliquer à la recherche de page Internet ou à la gestion du courrier électronique.

Bibliographie

SALTON, G. (1983). *Introduction to modern information retrieval*, New York, McGraw-Hill.

Proposition d'un classement des genres littéraires : une application de la théorie de l'énonciation

Sophie Piron
Université du Québec à Montréal

Introduction

Le but de mon étude est de proposer un classement des genres littéraires qui se situe à un macro-niveau d'analyse, sur la base de critères énonciatifs. Cet article clarifie d'abord le micro-niveau d'analyse, c'est-à-dire les genres littéraires. La seconde partie, consacrée au macro-niveau d'analyse, s'attache à la généricité en tant qu'attitude d'écriture, fonction textuelle. Dans ce travail, nous considérons que la problématique des genres littéraires relève de la typologie des discours car les textes littéraires sont des expressions particulières du discours (Todorov, 1987).

Prémisses théoriques : la répartition des genres littéraires

Pour proposer un macro-classement, il fallait d'abord étudier le genre en tant que classification historique. Les différentes formes d'écriture sont éminemment complexes, mais il était essentiel de les cerner par des définitions consensuelles qui pouvaient être appliquées aux textes modernes, c'est-à-dire essentiellement aux œuvres des XIX^{e} et XX^{e} siècles. L'étape préparatoire consistait donc à redéfinir de façon opérationnelle les différents genres littéraires.

Théâtre

Dans ce genre, chaque personnage prend tour à tour la parole.

Poésie

La poésie est reconnaissable par sa forme. En général, elle présente une disposition en vers réguliers ou libres qui forme une structure en lignes directement repérable. Avec la dislocation du vers, les poésies ont pu se transformer en textes brefs, mais elles demeurent reconnaissables par un aspect de leur forme, qui est la brièveté. En effet, il s'agit de textes courts (en prose) dont la longueur excède rarement une page.

Roman

Un roman est un texte assez long, en prose. Le trait le plus définitoire est que le roman est une fiction unitaire (Butor, 1992) dans le sens où il existe une unité au sein de l'œuvre à des niveaux divers (chronologie, personnages, etc.). Par ailleurs, ce que Butor désigne comme le contenant est parfois aussi peu fictif que narratif : des ouvrages comme ceux d'André Baillon se basent sur des faits réels ; d'autres, qui appartiennent au courant du Nouveau Roman, sont plus descriptifs que narratifs.

Conte et nouvelle

Le conte et la nouvelle sont tous deux des genres en prose relativement brefs. Tantôt volontairement confondus, tantôt mal différenciés, ils constituent les formes les plus ambiguës des typologies et histoires littéraires. Pour les définir, il faut se pencher sur leurs tactiques pour mener le récit au niveau de ses extrêmes (début et fin de l'histoire).

Le conte commence par sécréter son monde propre, par poser ses personnages et son intrigue[1]. Au contraire, une nouvelle aura tendance à débuter *ex abrupto*, à la façon d'un roman, ou posera d'emblée le cadre de son intrigue. Ceci permet de revenir à la situation antérieure qui est à son origine. À noter qu'il s'agit la plupart du temps d'un cadre plus réel – qui ne verse pourtant pas dans le réalisme le plus absolu – et non d'un monde délibérément fictionnel comme celui du conte.

En ce qui concerne la fin de ces récits brefs, le conte amène son intrigue à un bouclage complet (histoire « de a à z ») qui peut ajouter à la fin événementielle une morale, une réflexion, voire la contenir implicitement. La nouvelle boucle également son récit mais par un écho aux phrases, voire aux mots liminaires. Ainsi, la nouvelle présente une fiction unique. Ces limites imposent un seul chronotope.

Un autre critère de différentiation entre les deux genres concerne l'onomastique. Les personnages d'un conte sont sans véritable épaisseur psychologique et s'opposent facilement en couples d'actants. Pour cette raison, le conte emploie de préférence des groupes nominaux généraux pour désigner ses personnages[2]. La nouvelle use également de substituts onomastiques, mais de manière à davantage subjectiviser le récit puisque la psychologie y est plus développée. Ainsi, pour les actants, les prénoms et les pronoms personnels sont plus fréquents que dans un conte.

---

[1] Il faut aussi signaler que le conte peut également être identifiable par ses sous-types thématiques : le conte fantastique, le conte merveilleux, le conte de fées, le conte d'animaux et le conte à rire.

[2] « Le vieillard », « l'homme au ventre rebondi », « l'homme aux besicles ». Ces quelques exemples sont issus d'un conte qui s'intitule *Les trois médecins* (Enclin, 1915).

Autobiographie et mémoires

Ces deux genres sont très proches l'un de l'autre. Leur définition est une combinatoire de traits. Thématiques d'abord : l'histoire d'une existence, quelques épisodes d'une vie où l'accent est mis soit sur l'individualité – dans l'autobiographie[3] – soit sur des événements collectifs – dans les mémoires[4]. Des traits énonciatifs ensuite : narration autodiégétique rétrospective, qui est le plus souvent écrite à la première personne.

La définition de ces deux genres repose finalement sur « un contrat d'identité scellé par le nom propre » (Lejeune, 1975), reliant ainsi le narrateur au héros, et celui-ci à l'auteur.

Récit de voyage

Ce genre se caractérise lui aussi par une narration à la première personne. Sa thématique particulière, qui est un mélange de souvenirs et de discours cognitif (descriptions du pays visité en termes politiques, géographiques, etc.), le place à la charnière de l'autobiographie et de l'essai, affirmant par conséquent son individualité et posant son autonomie.

Essai

Son hétérogénéité thématique le rend difficilement définissable et particulièrement mal défini. D'ailleurs, certains le jugent extérieur au domaine littéraire. Cette place subalterne est en partie due à l'éventail des subdivisions dont le genre fait l'objet. Toutefois, malgré ses pratiques diverses, il connaît une certaine unité : la marque essentielle est l'expression d'une subjectivité plus ou moins fortement affirmée par l'auteur. Elle se déploie selon une construction idéelle argumentée avec plus ou moins d'intensité.

---

[3] « Biographie d'un auteur faite par lui-même » (*Petit Robert*, 1991).

[4] « Relation écrite qu'une personne fait des événements auxquels elle a participé ou dont elle a été témoin » (*Petit Robert*, 1991).

Chroniques

Les chroniques se présentent sous la forme d'une suite d'articles, de courts essais (variés ou sur un même thème). Le genre se manifeste par sa condensation par rapport à l'essai.

Aphorismes et maximes

L'aphorisme est proche du genre poétique par son travail de création langagière. C'est cet aspect poétique dominant qui définit sa particularité générique par rapport aux maximes. La pensée aphoristique est « prélogique » car elle évite tout enjeu délibérément philosophique (Lejeune et Voisin, 1988). Au contraire, les maximes restent fondamentalement réflexives. Elles se dévoilent explicitement comme pensée.

Le classement des genres littéraires qui vient d'être posé comprend ainsi 12 catégories.

TYPOLOGIE DES DISCOURS LITTÉRAIRES

Les systèmes antiques

Les systèmes antiques de classement tenaient compte de trois types de critères : la façon d'imiter (un critère philosophique), la présentation modale (un critère énonciatif) et l'histoire (un critère thématique). Chaque critère avait un statut propre, radicalement différent des autres, permettant de cette manière des définitions génériques à « géométrie variable » (Genette, 1986 : 105). Nous ne nous attachons ici qu'au mode d'énonciation.

Aristote proposait un système binaire[5] qui distingue deux formes de récit : le récit épique (dans lequel l'énonciation relève d'un narrateur) et le poème dramatique (dans lequel l'énonciation est le fait des personnages). Un autre système, proposé par Platon, distinguait trois formes de récit[6] : purement narratif (énonciation réservée au poète), mimétique (énonciation réservée aux personnages) et mixte (énonciation alternée entre le poète et les personnages). Au cours des siècles, la dyade aristotélicienne a été transformée en triade par l'ajout de la catégorie mixte de Platon.

Triade platonicienne	narratif	mixte	dramatique
Dyade aristotélicienne	épique		dramatique
Triade aristotélicienne : Compromis moderne	lyrique	épique	dramatique
Interprétation modale	Énonciation réservée au poète	Énonciation alternée entre le poète et les personnages	Énonciation réservée aux personnages

D'un point de vue énonciatif, les trois membres de la triade peuvent être entendus comme des attitudes fondamentales d'écriture ou de mise en forme[7]. Contrairement aux systèmes antiques qui étaient construits comme des grilles, ces trois catégories ont acquis un statut individuel. En suivant Genette (1986), nous considérons le lyrique, l'épique et le dramatique comme des archigenres, signifiant par là que ce sont des catégories plus générales qui surplombent les

[5] Et non un système ternaire, contrairement à ce que dit la tradition. D'ailleurs, ce système a été développé dans la *Poétique* et non dans la *Rhétorique*.

[6] Il est à noter que ce système ne s'intéressait qu'à l'imitation en vers.

[7] Le Romantisme et le Post-Romantisme n'ont plus envisagé ces catégories comme des modes d'expression, mais comme des genres allant jusqu'à comporter un élément thématique.

genres[8]. Généralement, on considère que le roman appartient à l'archigenre épique ; la poésie à l'archigenre lyrique et le théâtre à l'archigenre dramatique :

Archigenre épique	Archigenre lyrique	Archigenre dramatique
↓	↓	↓
Genre : roman	Genre : poésie	Genre : théâtre

L'éventail des genres que nous avons préalablement établi s'intègre difficilement dans cette triade. C'est la raison pour laquelle nous proposons une typologie plus étendue dont les critères énonciatifs sont définis systématiquement.

CRITÈRES POUR UNE TYPOLOGIE DES DISCOURS LITTÉRAIRES

Premier critère : les macro-actes de langage

Les genres littéraires, une fois considérés comme des discours, sont classés dans des groupes auxquels correspondent des actes de langage. Les actes définis par Austin et Searle ne peuvent être transposés ici car ils ont le désavantage tantôt de définir des sous-genres[9], tantôt de rapprocher des genres[10].

---

[8] Gérard Genette considère cependant que ces catégories comportent toujours un élément thématique, aussi vague soit-il. Nous pensons que la littérature moderne a bouleversé ces frontières thématiques et qu'il est donc plus pertinent de se baser sur des critères uniquement énonciatifs.

[9] La théorie d'Austin et Searle classe les discours en groupes auxquels correspondent des actes de langage. On a ainsi les actes assertifs (raconter, décrire, commenter...), promissifs (promettre), expressifs (exhorter, supplier, demander...), directifs (ordonner, réclamer, demander...) et déclaratifs (narrer, exposer...). L'inconvénient pour l'étude des discours littéraires est, par exemple, qu'un acte directif comme « ordonner » peut intervenir dans certains discours idéologiques (sous-genre de l'essai).

[10] Les actes qui consistent à interpeller (exhorter, supplier, demander...) sont utilisés dans la poésie et dans le théâtre.

Nous proposons donc des actes de langage (actes illocutoires) dont la portée plus spécifique permet de mieux délimiter les catégories du discours littéraire. De plus, la distribution est construite d'après la notion de dominante car un genre peut engager plusieurs actes de langage. C'est pourquoi nous parlerons de « macro-actes de langage ».

Le roman, le conte et la nouvelle ont comme macro-acte de langage celui de raconter car tous trois narrent une histoire. Le récit de voyage, l'autobiographie et les mémoires font de même, si ce n'est qu'ils content des faits qui sont en relation avec une expérience vécue par l'auteur. Le macro-acte mis en œuvre consiste ainsi à se raconter. Le théâtre est le seul genre dans lequel les personnages prennent la parole tout au long de l'histoire. Pour celui-ci, l'acte illocutoire dominant est dialoguer. On peut nommer l'acte exalter pour la poésie et les aphorismes car ces catégories littéraires cherchent à atteindre un haut degré d'intensité expressive. Enfin, l'acte de langage qui domine dans l'essai, les chroniques et les maximes est l'acte qui revient à exposer puisque des idées sont exposées, présentées (avec une certaine neutralité ou bien en défendant ou attaquant une position particulière).

Deuxième critère : la forme d'énonciation

Une répartition des genres littéraires simplement basée sur l'opposition traditionnelle prose / poésie est un critère non discriminant. Au contraire, la forme d'énonciation permet de dépasser cette opposition formelle superficielle (micro-niveau des phrases) car elle caractérise le macro-niveau, celui du discours. D'ailleurs, la forme d'énonciation prolonge le choix d'un macro-acte de langage.

Dans de nombreux genres (roman, conte, nouvelle, récit de voyage, autobiographie et mémoires), le récit est essentiellement présenté

par un narrateur, qu'il soit omniscient ou non. Ces récits sont des agencements d'événements. Le fait que ceux-ci sont tantôt fictifs, tantôt réels, n'a pas d'incidence sur la forme d'énonciation qui reste narrative[11]. Pour ce qui est du théâtre, sa forme est dialogique[12]: l'histoire se déroule en dialogues, démultipliant ainsi les narrateurs / énonciateurs. La poésie et les aphorismes se manifestent par leur travail sur la langue. Ces genres peuvent démotiver les relations grammaticales et lexicales de l'usage quotidien pour donner un sens pluriel et interne au texte. La forme d'énonciation est tropique[13]. Enfin, les textes que l'on retrouve dans les essais, les chroniques et les maximes se présentent quant à eux sous forme d'exposés argumentatifs. Leur énonciation est donc expositive[14].

Troisième critère : la finalité d'énonciation

La finalité du théâtre, des romans, contes et nouvelles est la représentation[15] car ils posent une réalité mimétique. Ils se présentent en conséquence à l'écart du système énonciatif car, même s'il existe un Je-Origine, celui-ci est fictif et attire avec lui dans la fiction, le système énonciatif créé (Hamburger, 1986).

La poésie et les aphorismes s'intègrent dans le système énonciatif car ce sont des énoncés d'un sujet d'énonciation qui s'identifient avec un Je-Origine réel (le poète), mais ces genres se situent à l'écart de la réalité car ils déploient une réalité subjective qui attire le pôle-objet vers le pôle-sujet. Leur finalité réside dans cette présentation.

---

11 Ces genres sont tous écrits en prose.

12 Le théâtre peut être écrit en vers ou en prose.

13 Ces deux genres sont tantôt en prose, tantôt en vers.

14 Ces genres sont écrits en prose.

15 Les critères de représentation et de présentation sont issus de Todorov (1987).

L'autobiographie, les mémoires et les récits de voyage posent l'énoncé dans la réalité. Celle-ci est parfois tronquée, transformée, mais jamais fictionnalisée. La finalité consiste à reproduire une réalité qui est celle du pôle-sujet. Les essais, chroniques et maximes font également référence à des arguments, une réalité qui n'est pas fictionnalisée. Ici aussi, il s'agit d'une énonciation référentielle. Toutefois, sa particularité réside dans le glissement vers le pôle-sujet car l'auteur présente sa vision des faits.

Quatrième critère : la fonction de communication / énonciation

Jakobson (1978) a proposé six fonctions de communication qui correspondent aux six éléments constitutifs du message linguistique[16]. Ceci peut être appliqué de façon systématique aux discours littéraires.

Le roman, le conte et la nouvelle narrent des faits. Ils s'attachent à un contexte d'événements (fictifs). Leur fonction est donc référentielle puisqu'elle est centrée sur le contexte. Les genres qui narrent des événements vécus par l'énonciateur (récit de voyage, autobiographie, mémoires) se centrent sur le destinateur, autrement dit sur la fonction émotive. Le genre qui s'attache par excellence au moyen de contact (fonction phatique) utilisé entre les personnages est le théâtre. La poésie et les aphorismes insistent sur d'autres fonctions qui sont soit la fonction métalinguistique (le code), soit la fonction émotive (le destinateur). En effet, ces genres s'attachent à l'onirisme des mots, à la pluralité de leur signifié et / ou expriment des sentiments personnels. Finalement, l'essai, les chroniques et les maximes sont des genres qui sont le plus souvent tournés vers les

16 Jakobson lui-même a spécifié la triade classique au moyen de sa théorie : la fonction émotive pour le lyrisme, la fonction référentielle pour l'épique et la fonction conative pour le dramatique (parce qu'il serait centré sur l'effet à produire). Dans ce dernier cas, Jakobson semble davantage insister sur la représentation théâtrale que sur le discours littéraire.

idées qu'ils proposent, ils peuvent aussi tenter de rallier le lecteur à un point de vue, le convaincre... C'est pourquoi ces types de discours mettent en exergue le message (fonction poétique) ou le destinataire (fonction conative).

Typologie des discours littéraires

Les différents critères élaborés au cours de la section précédente convergent vers un classement des discours littéraires en cinq archigenres : épique (roman, conte, nouvelle), mnémique[17] (récit de voyage, autobiographie, mémoires), dramatique (théâtre), lyrique (poésie, aphorismes) et logique[18] (essai, chroniques, maximes).

Notre taxonomie des actes d'écriture fondamentaux montre à quel point la théorie de l'énonciation contribue à établir des modèles d'analyse. Ainsi, nous avons appliqué notre typologie à la *Bibliographie des écrivains français de Belgique*, et ceci nous a permis de démontrer que les écrivains belges du début du XXe siècle ont réparti leurs publications entre de nombreux archigenres, contrairement aux écrivains postérieurs qui, tout en publiant dans de nombreux genres, ne présentent pas un si haut degré de variation archigénérique (Piron, 2000).

---

17 Nous avons donné à cette catégorie une étymologie grecque : *μνημη* signifie « le souvenir ».

18 Nous avons ici aussi formé un terme à partir d'une étymologie grecque : *λογος* qui signifie « la parole, le propos ».

Tableau synoptique

ARCHIGENRE	GENRE	MACRO-ACTE DE LANGAGE	FORME D'ÉNONCIATION	FINALITÉ D'ÉNONCIATION	FONCTION DE COMMUNICATION
Épique	Roman Conte Nouvelle	Raconter	Narrative	Représentation Créer une mimesis de réalité	Fonction référentielle
Mnémique	Récit de voyage Autobiographie Mémoires	Se raconter	Narrative	Référence Reproduire la réalité du pôle-sujet	Fonction émotive
Dramatique	Théâtre	Dialoguer	Dialogique	Représentation	Fonction phatique
Lyrique	Poésie Aphorismes	Exalter	Tropique	Présentation	Fonction métalinguistique et / ou émotive
Logique	Essai Chroniques Maximes	Présenter	Expositive	Référence Glissement vers le pôle-sujet	Fonction poétique et / ou conative

46

Bibliographie

AUSTIN, J.L. (1970), *Quand dire, c'est faire*, Paris, Seuil.

BAUDELLE, Y. (1991), « Nouvelle et noms propres 1920-1959 », *La nouvelle (définitions, transformations)*. Colloque, Textes recueillis par B. ALLUIN et F. SUARD, Lille, Presses Universitaires de Lille, UL3.

BUTOR, M. (1992), *Essais sur le roman*, Paris, Gallimard.

ENCLIN, V. (1915), *Les histoires de Peter*, Renaix, Leherte-Courtin éditeurs.

GENETTE, G. (1986), « Introduction à l'architexte », *Théorie des genres*, Paris, Seuil, p. 89-159.

HAMBURGER, K. (1986), *La logique des genres littéraires*, Paris, Seuil.

JAKOBSON, R. (1978), *Essais de linguistique générale. 1. Les fondations du langage*, Paris, Minuit.

LEJEUNE, C., et M. VOISIN (1988), « Entretien de Marcel Voisin avec Claire Lejeune », dans D. BLAMPAIN, A. NYSENHOLC, J. RUBES, M. TROUSSON et M. VOISIN, *L'essai en Belgique romane*, numéro spécial d'études littéraires, 21, 2, automne 1988, Québec, Université Laval.

LEJEUNE, P. (1975), *Le pacte autobiographique*, Paris, Seuil.

PIRON, S. (2000), « La diversité générique a-t-elle caractérisé la littérature belge ? », dans M. RAJMAN et J.-C. CHAPPELIER (éd.), *JADT 2000 : Actes des cinquième journées internationales d'analyse statistique des données textuelles, 9-11 mars 2000*, EPFL, Lausanne, 2, p. 467-475.

SEARLE, J. (1972), *Les actes de langage*, Paris, Hermann.

SIMONSEN, M. (1987), « Le Conte », dans J.-P. BEAUMARCHAIS, D. COUTY et A. REY, *Dictionnaire des littératures de langue française,* Paris, Bordas.
TODOROV, T. (1987), *Notion de littérature et autres essais*, Paris, Seuil.

VIËTOR, K. (1986), « L'histoire des genres littéraires », *Théorie des genres*, Paris, Seuil.

Impact du contact des langues en milieu scolaire à l'île Maurice

Karina Chavry
avec la collaboration de *Denise Deshaies*
Université Laval

Introduction

Les phénomènes linguistiques liés au contact des langues, tels que les emprunts, les calques et les interférences, sont observables dans toute communauté plurilingue et pluriculturelle, comme c'est le cas de l'île Maurice. Ces phénomènes peuvent être observés à l'oral aussi bien qu'à l'écrit, et ce, dans toutes les sphères sociales. Lorsqu'ils sont présents dans le secteur éducatif, ou plus particulièrement dans des devoirs écrits d'élèves du cycle secondaire en français, desquels est constitué notre corpus, les enjeux de leur présence sont d'autant plus importants qu'ils peuvent être la cause d'échecs ou de résultats scolaires insatisfaisants. Par conséquent, l'identification de ces phénomènes linguistiques et la reconnaissance du fait que les manifestations du contact des langues est une réalité normale pourraient contribuer à modifier la pédagogie du français dans cette communauté, puisque le français y est enseigné comme langue maternelle, alors qu'il devrait être enseigné comme une langue seconde.

La problématique

Les phénomènes linguistiques liés aux langues en contact (emprunts, calques, alternances de codes et interférences) ont été intensivement étudiés, et ce, dans des cadres théoriques variés. Beaucoup d'entre eux ont été analysés à partir de corpus oraux (Poplack, 1980 ; Poplack et Sankoff, 1984 ; Flikeid, 1989), mais peu ont été observés à l'écrit, comme c'est le cas de notre étude.

De même, peu de travaux ont porté sur le décalage entre la réalité linguistique et l'enseignement des langues dans une optique sociolinguistique, encore moins au niveau du cycle secondaire, dans le contexte mauricien. Même si beaucoup de travaux ont été faits sur les problèmes liés à l'enseignement du français au cycle primaire à l'île Maurice (Tirvassen, 1994), peu de recherches ont eu pour objectif de se consacrer à l'étude des effets linguistiques dus aux langues en contact dans les écrits d'élèves à tous les niveaux d'un lycée en particulier.

Selon les chiffres proposés par Baggioni et Robillard (1990), il n'y a qu'entre 2% et 5% de la population mauricienne qui a le français comme L1 (et il s'agit ici du français mauricien, et non du français standard), entre 55% et 60% l'ont en L2 et 5% en L3. Ces chiffres, confirmés par Hookoomsing (1997), indiquent qu'au moins 95% des élèves qui étudient le français à l'île Maurice n'ont pas le français comme L1, alors qu'ils l'apprennent comme si c'était le cas, notamment au cycle secondaire. Il semble donc, d'après ces chiffres, que le français « langue seconde » devrait être enseigné à ce niveau.

Dans cet article, nous ferons état des résultats de nos recherches en les présentant selon trois des variables utilisées : la nature grammaticale des données, leur langue source et l'appartenance ethnique des informatrices.

Phénomènes liés aux langues en contact

Les phénomènes linguistiques liés aux langues en contact sont plus connus sous les termes d'emprunt, d'alternance de codes, d'hypercorrection et d'interférence et ont fait l'objet de plusieurs études, notamment dans un cadre sociolinguistique, comme ce sera le cas ici.

L'alternance de codes

Poplack, (1980 : 583), définit l'alternance de codes comme « *the alternation of two languages within a single discourse, sentence or constituent* ». Plus tard, dans le cadre de son étude sur le français d'Ottawa et de Hull, elle ajoute que les alternances de codes sont des « énoncés composés de plusieurs mots qui ne sont pas intégrés au lexique, à la syntaxe et à la morphologie du français », comme c'est le cas dans un des exemples qu'elle donne « Tu sais, *I helped them* à comprendre le français pis à parler » (Poplack, 1990 : 134).

L'hypercorrection

Dubois *et al.* (1994 : 236) présentent l'hypercorrection comme étant « la faute ou l'écart phonétique, morphologique, lexical ou syntaxique, né de l'application d'une norme ou d'une règle là où il n'y a pas lieu, par le jeu de l'analogie des paradigmes ». Cependant, le *Nouveau Petit Robert 1993* met en perspective une dimension importante, particulièrement applicable à la situation sociolinguistique mauricienne. Il la définit comme étant une « reproduction fautive d'une forme linguistique produisant une forme supposée correcte et par extension, le fait de produire des formes linguistiques anormales ou fautives par souci de manifester une maîtrise du discours signalant un statut social valorisé ».

Les hypercorrections relevées dans notre corpus sont dues, pour la plupart, au contact entre le créole et le français. Et c'est le rapport diglossique qui existe entre ces deux langues qui fait que le créole se présente comme une langue aussi largement utilisée que stigmatisée. Des exemples d'hypercorrections sont « vétérinaire » qui devient *vétérineuse* et « friandises » qui devient *friandisses*.

L'emprunt

Selon Dubois *et al.* (1994 : 177-178), « il y a emprunt linguistique quand un parler A utilise et finit par intégrer une unité ou un trait linguistique qui existait précédemment dans un parler B (dit langue source) et que A ne possédait pas ; l'unité ou le trait emprunté sont eux-mêmes qualifiés d'emprunts. [...] L'intégration, selon qu'elle est plus ou moins complète, comporte des degrés divers... ».

Nous avons ainsi dégagé de notre corpus des emprunts lexicaux qui sont des importations directes de mots d'une langue A dans une langue B, par exemple *job*, et des emprunts sémantiques qui se résument en l'attribution du sens d'un mot d'une langue A à un mot d'une langue B, par exemple *réaliser*, utilisé dans le sens de « se rendre compte de ».

L'interférence

Dubois *et al.* (1994 : 252) présentent le phénomène linguistique de l'interférence comme existant « quand un sujet bilingue utilise dans une langue-cible A un trait phonétique, morphologique, lexical ou syntaxique caractéristique de la langue B ». Selon eux, « [...] l'interférence reste individuelle et involontaire... ».

De notre corpus, nous avons recensé six types d'interférences : les interférences sémantiques, morphologiques, phonologiques, lexicales, syntaxiques et graphiques.

Le corpus et l'échantillon

Notre corpus est constitué de rédactions écrites par des élèves du Collège de Lorette de Port-Louis qui a officiellement l'anglais comme langue d'enseignement. La même rédaction (« La personne que j'admire le plus ») a été donnée et faite par la majorité des

élèves choisies, et ce, indépendamment de leur niveau académique et de leur âge.

Pour établir un échantillonnage assez représentatif, les éléments suivants ont été pris en compte.

Premièrement, nous avons choisi une seule classe par niveau académique, les niveaux étant appelés, dans cette école, des Forms. La Form I est un niveau qui comprend des filles qui ont 12 ou 13 ans. À l'opposé, nous retrouvons la Form VI dont les élèves ont au maximum 20 ans. Il faut noter que la Form VI est faite en deux ans[1], la VI[1] et la VI[2]. Nous avons donc un total de 7 classes, une classe par Form, dont les élèves sont âgées de 12 à 19 ans. Les Forms I, II et III sont divisées en trois classes chacune : rouge, vert et bleu. Les classes choisies sont les Forms I Bleu, II Vert et III Bleu. En ce qui concerne les Forms IV et V, elles comprennent quatre classes chacune : A-D, celles retenues étant les Forms IV D et V D. Nous avons un total de 167 copies, dont la longueur moyenne est de 35 lignes pour les plus jeunes et de 55 pour les plus âgées.

Le choix de ces classes a principalement été fait selon le niveau général de la connaissance en français et de la disponibilité des élèves et de leur enseignant(e) de français. Par conséquent, pour les cinq premiers niveaux, ce sont les classes qui ont, en général, les résultats les plus faibles en français qui ont été sélectionnées, car la sélection de classes réputées pour être d'un bon niveau en français aurait peut-être rendu nulle notre recherche. En ce qui concerne les deux derniers niveaux, les classes étant organisées selon le choix de matières des élèves plutôt que selon les résultats académiques, nous avons choisi les classes avec lesquelles nous avions déjà travaillé.

[1] Les Forms I-V correspondent aux classes de Secondaire I-V du système québécois. Les Forms VI, première et deuxième années, correspondent aux deux années de CÉGEP.

Les résultats

Ainsi que mentionné plus haut, nos données sont issues de 167 copies et ont mis 310 phénomènes linguistiques liés aux langues en contact en évidence. Ces 310 phénomènes ont été produits par 114 filles, soit 68,3% des informatrices. Ils sont présentés plus bas selon trois critères d'analyse : la nature grammaticale et la langue source des phénomènes ainsi que l'appartenance ethnique des informatrices.

Le tableau 1 présente tous les phénomènes linguistiques à l'exception de l'alternance de codes, selon leur nature grammaticale, et en ordre décroissant.

Tableau 1 : Les résultats généraux selon la nature grammaticale des phénomènes

	Hypercorrection	Emprunt	Interférence	Total
Nom commun	5 (4,7%)	22 (20,7%)	79 (74,5%)	106 (34,3%)
Adjectif	8 (16,3%)	23 (46,9%)	18 (36,7%)	49 (15,8%)
Nom propre	0 (0,0%)	45 (97,8%)	1 (2,2%)	46 (14,8%)
Pronom personnel	0 (0,0%)	0 (0,0%)	36 (100%)	36 (11,6%)
Verbe	2 (6,5%)	4 (12,9%)	25 (80,6%)	31 (10,0%)
Syntagme / Proposition	0 (0,0%)	0 (0,0%)	19 (100%)	19 (6,1%)
Pronom relatif	0 (0,0%)	0 (0,0%)	6 (100%)	6 (1,9%)
Adverbe	0 (0,0%)	0 (0,0%)	5 (100%)	5 (1,6%)
Préposition	0 (0,0%)	2 (40,0%)	3 (60%)	5 (1,6%)

Locutions	0 (0,0%)	0 (0,0%)	5 (100%)	5 (1,6%)
Article	0 (0,0%)	0 (0,0%)	1 (100%)	1 (0,3%)
Total	15 (4,9%)	96 (31,0%)	198 (64,0%)	309

Tel que nous pouvons le constater à partir de ce tableau, les noms communs arrivent largement en tête, suivis de loin par les adjectifs et les noms propres. Les verbes n'arrivent qu'en cinquième position, après les pronoms personnels, lesquels sont tous impliqués dans l'interférence syntaxique. Nous voyons donc que les catégories ouvertes, soit les noms, les verbes et les adjectifs, dont parle Poplack (1990), sont bel et bien les plus représentées dans notre corpus écrit.

Si l'on regarde les résultats de certaines études sociolinguistiques qui ont été menées dans des communautés francophones en Ontario, en Nouvelle-Écosse et au Nouveau-Brunswick sur les emprunts (Flikeid, 1989 ; Picard, 1996), des résultats presque similaires ont été obtenus quant à la nature des mots empruntés. Il est intéressant de noter ici que, contrairement à ces études faites à partir de corpus oraux, le nôtre présente plus d'emprunts adjectivaux que d'emprunts verbaux, les verbes se retrouvant plus dans les interférences, donc dans un type de phénomène considéré comme étant inconscient et involontaire. Peut-être est-ce parce que les verbes, dans le contexte du sujet de la rédaction proposée, sont moins à même d'être empruntés, ce qui est loin d'être le cas des noms et des adjectifs, ou alors parce que l'on est à l'écrit et que les élèves font plus attention aux verbes qu'elles utilisent, ces mots étant souvent le point faible des apprenantes.

Le prochain tableau est celui concernant la langue source des phénomènes.

Tableau 2 : Les résultats généraux selon la langue source des phénomènes

	Créole	Anglais	Créole ou Anglais	Français mauricien	Autre
Alternance de codes	0	1 (100,0%)	0	0	0
Hypercorrection	14 (93,3%)	1 (6,7%)	0	0	0
Emprunts : noms propres	0	42 (93,3%)	0	1 (2,2%)	2 (4,4%)
Emprunts : autres	21 (41,2%)	30 (58,8%)	0	0	0
Interférences sémantiques et morphologiques	10 (100,0%)	0	0	0	0
Interférences phonologiques	17 (100,0%)	0	0	0	0
Interférences lexicales	9 (90,0%)	1 (10,0%)	0	0	0
Interférences syntaxiques	15 (23,8%)	7 (11,1%)	40 (63,5%)	1 (1,6%)	0
Interférences graphiques	0	98 (100,0%)	0	0	0
Total	86 (27,8%)	180 (58,1%)	40 (12,9%)	2 (0,6%)	2 (0,6%)

Le tableau 2 démontre que le créole influence de manière certaine le français à l'écrit dans notre corpus, surtout en ce qui concerne les éléments phonologiques, morphologiques, sémantiques et lexicaux du discours. L'importance du créole dans les interactions quotidiennes des informatrices est ici mise en évidence par le rôle que joue cette langue dans les écrits des élèves aux plans phonologique, morphologique, lexical et sémantique. Nous notons que cette langue a aussi une importance relative au niveau de la syntaxe, d'autant plus qu'il est beaucoup plus probable que les

élèves soient plus influencées par le créole que par l'anglais lors de la construction de leurs phrases. L'anglais se retrouve donc dans des catégories bien spécifiques qui sont surtout les emprunts, particulièrement les noms propres, où l'anglais intervient dans 93,3% des cas, et les interférences graphiques, où l'anglais est impliqué dans 100% des cas, les interférences graphiques représentant 49,5% l'ensemble des interférences. Le fait que les interférences graphiques proviennent à 100% de l'anglais n'est pas étonnant puisque les élèves sont scolarisées en anglais (tous les travaux sont écrits en anglais, sauf ceux du cours de français et de littérature française) et non en créole qui est une langue qu'elles n'écrivent pas, du moins à l'école. En ce qui concerne les noms propres empruntés, qui le sont surtout de l'anglais, nous pensons que cela s'explique par le fait que cette langue étant la langue administrative du pays, se retrouve dans toutes les appellations officielles et dans tout ce qui touche la vie scolaire. Ces noms sont donc des emprunts obligatoires, d'autant plus qu'ils sont des noms propres. En somme, il est intéressant de voir que le rôle de chacune des langues dans la vie quotidienne de l'Île Maurice se traduit par l'influence respective que celles-ci ont sur les divers plans des discours écrits des informatrices.

Dans le tableau 3 suivant, nous avons une perspective générale des résultats par rapport à l'appartenance ethnique des 114 filles ayant produit des phénomènes linguistiques liés aux langues en contact.

Tableau 3 : Les résultats généraux selon l'appartenance ethnique des informatrices

	Chinoise	Créole	Hindoue	Musulmane	Tamoule
Alternance de codes	0	0	0	1	0
Hypercorrection	1	12	1	1	0
Emprunt	31	41	5	12	7
Interférence	41	81	17	43	16

Nombre de phénomènes par ethnie	73	134	23	57	23
Nombre d'informatrices concernées	25	54	9	20	6
Moyenne	2,92	2,48	2,55	2,85	3,83
Rang	2	5	4	3	1
Total d'informatrices	37	84	15	23	8
Moyenne par informatrice sur le total du groupe	1,97	1,59	1,53	2,48	2,87
Rang	3	4	5	2	1
% d'informatrices concernées	67,6%	64,3%	60%	86,9%	75%

Ce tableau permet de constater que si 60% et plus des informatrices ont produit des phénomènes liés aux langues en contact, elles sont plus nombreuses parmi les Musulmanes et les Tamoules. Ces deux groupes sont également ceux pour lesquels on observe les moyennes sur le total du groupe les plus élevées. Par ailleurs, si l'on regarde la moyenne par élève concernée, seules les Tamoules se démarquent des autres groupes, les autres ayant des moyennes assez semblables. Le faible nombre d'informatrices dans ce groupe peut expliquer ce résultat. Cependant, nous notons que les différences quant à l'appartenance ethnique sont quand même relativement négligeables.

Nous avancerons que la croyance mythique à Maurice selon laquelle les ethnies sont quasi génétiquement plus portées vers une langue plutôt que vers une autre n'a aucun fondement linguistique,

ainsi que nous pouvons le constater ici. Les choix linguistiques sont certes historiques et politiques, mais en aucun cas génétiques.

Conclusion

Ainsi, il existe bel et bien un écart entre la pratique linguistique des élèves et le français standard qui leur est enseigné étant donné que nous avons recensé 310 phénomènes linguistiques liés aux langues en contact. Il apparaît donc clairement que les deux langues « supra-communautaires » que sont l'anglais et le créole influencent la troisième qui est le français. Le fait que le français et le créole influencent l'anglais nous paraît fort possible étant donné que nous avons trouvé une interférence graphique du français sur l'anglais au sein de la seule alternance de codes de notre corpus. Cela pourrait certainement faire l'objet d'une autre étude.

Nous tenons cependant à souligner qu'à partir d'un total de 47 000 mots, notre corpus n'a révélé que 310 phénomènes liés aux contacts des langues, et, qu'en plus, il s'agit ici de filles considérées comme étant les moins fortes en français de la Form I à la Form V. Ceci indique donc que l'influence de l'anglais et du créole sur le français est assez minime. Toutefois, nous pensons que cela n'empêche pas l'apport de certaines améliorations au système d'enseignement du français, bien au contraire, il souligne son importance.

Bibliographie

BAGGIONI, D., et D. de ROBILLARD (1990), *Île Maurice : une francophonie paradoxale*, Paris, L'Harmattan.

DUBOIS. J. (dir) (1994), *Dictionnaire de linguistique et des sciences du langage*, Paris, Larousse.

FLIKEID, K. (1989), « "Moitié anglais, moitié français ?" Emprunts et alternance de langues dans les communautés acadiennes de la Nouvelle-Écosse », *Revue québécoise de linguistique théorique et appliquée*, 8, 2, avril 1989, p. 177-228.

HOOKOOMSING, V. (1997), « Pluralisme linguistique et développement à l'île Maurice : quelques réflexions à partir du recensement de 1990 », dans M.C HAZAEL-MASSIEUX et D. de ROBILLARD (éd.), *Contacts de langues, contacts de cultures, créolisation*, Paris et Montréal, L'Harmattan, p. 387-409.

PICARD, A. (1996), *Faits de variation dans le parler d'adolescents du Nouveau-Brunswick : français acadien et contact avec l'anglais*. Mémoire de maîtrise, Québec, Université Laval.

POPLACK, S. (1980), « "Sometimes I'll start a sentence in Spanish *y termino en español* " : toward a typology of code-switching », *Linguistics*, 18, p. 581-618.

POPLACK, S. (1990), « Statut de langue et accommodation langagière le long d'une frontière linguistique », dans R. MOUGEON et É. BÉNIAK (dir.), *Le français canadien parlé hors Québec*, Québec, Presse de l'Université Laval, p. 127-151.

POPLACK, S., et D. SANKOFF (1984), « Borrowing : the synchrony of integration », *Linguistics*, 22, p. 99-135.

TIRVASSEN, R. (1994), *Communication et développement : le cas de Maurice*, Aix-en-Provence, Didier Érudition.

Ceci n'est pas une pipe

François Parent
Université Laval

> *Pour certaines personnes la langue, ramenée à son principe essentiel, est une nomenclature, c'est-à-dire une liste de termes correspondant à autant de choses [...] cette conception est critiquable à bien des égards. Elle laisse supposer que le lien qui unit un nom à une chose est une opération toute simple, ce qui est bien loin d'être vrai* (Saussure, 1975 : 97).

Introduction

Historiquement on considérait que le nom renvoie à la chose qu'il désigne. Encore aujourd'hui, faute de discerner sa valeur immanente propre, la considération de la valeur sémantique d'un mot se limite fréquemment à sa valeur référentielle, celle que nous propose, par exemple, la définition du dictionnaire. En regard de leur besoin de décrire le monde et de communiquer, les locuteurs n'ont pas à se préoccuper de la valeur conceptuelle que la langue tient à leur disposition, mais uniquement de la valeur référentielle qu'elle prend dans l'usage. Cette tendance naturelle à ignorer la valeur immanente du concept nuit à une explication adéquate du phénomène de la polysémie et de celui de la synonymie. Par exemple, on admet difficilement que « pomme » dans les séquences suivantes :

> une *pomme* de terre rouge ;
> la *pomme* verte est pourrie

soit un seul et même concept, dont le référent change en fonction du contexte d'emploi. Certains considèrent même que la séquence ***pomme de terre*** serait une sorte de mot complexe puisqu'elle peut avoir le même référent que *patate*. Mais cette

synonymie est un fait de concordance référentielle qui n'affecte pas la valeur immanente du concept « pomme » dans cette séquence. Cette communication vise à reconnaître le rapport qui s'institue entre la valeur immanente et la valeur référentielle des concepts afin d'expliquer concrètement les phénomènes de la polysémie et de la synonymie dans le cadre de l'approche théorique de la sémantique grammaticale.

Du mot en tant que reflet de la réalité

Historiquement on considérait que le nom correspond à la chose qu'il désigne et qu'il imite la nature : il en serait en quelque sorte le reflet. Pour les Grecs par exemple, le nom appartenait à la chose, puisque tout était d'inspiration divine. Encore aujourd'hui, on considère en certains cas que le nom renvoie à la chose. Un auteur comme Van Vlasselaer considère même que :

> *Le mot, même à l'époque contemporaine, retient son attrait particulier dans le cas du « tabou » qui, bien qu'au départ de nature conventionnelle, tend à faire confondre le mot à tel point avec la chose, que l'un évoque l'autre avec une telle intimité qu'une maladie renaît avec son appellation, qu'un animal dangereux apparaît avec son nom, qu'un dieu foudroie avec son évocation le malheureux mortel qui hasardait son nom* (1973 : 23-24).

Or, aujourd'hui, on admet généralement que le mot n'est pas la chose, pas plus que la reproduction picturale, ce que voulait mettre en évidence Magritte avec son œuvre *Ceci n'est pas une pipe*, soulignant le fait que l'image d'une pipe ne constitue pas la pipe elle-même : il s'agit d'une illustration seulement. Ce point de vue rejoint celui de Benjamin Lee Whorf : « À quoi donc supposons-nous que notre moi conscient se trouve confronté lorsque nous pensons à ce rosier ? Nous croyons probablement qu'il a affaire à une "image mentale" qui n'est pas le rosier mais une substitution mentale » (1971 : 101).

En pratique, on confond pourtant fréquemment la représentation mentale figurative qu'est le référent avec la chose qui existe dans le monde physique. En effet, le mot en tant que signifié est un objet mental, non un objet de l'univers externe à la pensée. Il n'est donc jamais la chose à laquelle il renvoie dans l'usage : tout comme « la carte n'est pas le territoire » (Schaff, 1969 : 265) mais une représentation du territoire. Si cette distinction n'apparaît pas tout à fait évidente dans le cas de noms comme *chaise* et *chat* – qui renvoient à des objets du monde physique –, des termes comme *licorne*, *griffon* et *dragon*, soulignent bien cette distinction entre le référent et l'objet du monde extra-mental. En effet, ces noms, qui relèvent de la mythologie, ont dans l'usage un référent qui ne correspond pas à une réalité physique mais à une réalité mythologique qui tient de l'imagination.

Concept et référence

L'opposition du concept et du référent est une autre distinction à laquelle les spécialistes en linguistique sont confrontés dans leur analyse. Souvent, on ne considère même pas le fait qu'un mot véhicule nécessairement ces deux valeurs.

Pourtant, Wittgenstein soulignait que, si l'on considère la langue comme une nomenclature, les mots gagnent peut-être en stabilité mais, lorsqu'on les met ensemble dans une phrase, on ne comprend pas comment ils pourraient servir à communiquer quelque chose de nouveau. En effet, le fait qu'un même mot se prête à diverses références implique qu'il puisse s'adapter à différents contextes d'emploi. En outre, son identité est liée à la stabilité de sa valeur. Ces deux points de vue ne sont pas contradictoires : ce sont deux pôles sémantiques complémentaires caractérisant les signifiés marqués par les mots.

Certains spécialistes commencent à cerner cette distinction, comme Victorri et Catherine Fuchs qui parlent d'un noyau de sens (1996 : 51) qui ne serait pas un sens à proprement parler, mais un « schéma organisateur, sorte de base à partir de laquelle

se construisent les différents sens de l'unité, y compris le sens premier [...] » (*Ibid.*). De son côté, Picoche, dans le cadre guillaumien, envisage un sens de l'unité appelé saisie plénière, (dans Victorri et Fuchs, 1996 : 51). Cette auteure illustre ce qui serait le processus de dérivation des autres emplois en expliquant différents emplois du terme *toucher* d'après son emploi plénier dans la phrase *Jean touche l'épaule de Paul.* C'est l'emploi qu'elle considère le plus riche sémantiquement, vu sa correspondance au fait que :

> *un sujet humain A éprouve une sensation tactile en mettant intentionnellement l'épiderme de sa main en contact avec l'épiderme d'un sujet humain B qui, à son tour, en éprouve une sensation tactile, plus ou moins clairement porteuse de l'information qu'a voulu y mettre A* (*Ibid*).

De cet emploi plénier dériveraient trois sens principaux par des cinétismes qui produisent chacun des autres sens ou saisies. Par exemple, le cinétisme I, défini par le fait que l'humain touche un inanimé, passerait par trois saisies allant du concret vers l'abstrait :

> 1re saisie : Je touche le fer pour voir s'il est chaud (*Ibid.*).
> 2e saisie : Ne pas toucher aux objets exposés (*Ibid.*).
> 3e saisie : Toucher à tous les sujets d'actualité (*Ibid.*).

Tandis que le cinétisme II, défini par le fait qu'un inanimé touche un animé, donne lieu aux deux saisies suivantes :

> 1re saisie : La balle touche Paul (*Ibid.*).
> 2e saisie : Ce reproche touche Paul (*Ibid.*).

et ainsi de suite pour le troisième cinétisme.

Cependant, comme le remarquent Victorri et Fuchs dans leur ouvrage, ce schéma pose problème parce qu'il ne permet pas encore de rendre compte du sens figuré de certaines unités, comme par exemple *table* et *arbre* dans *table de multiplication* ou *arbre généalogique*. À ce propos, ces auteurs signalent

l'importance de se dégager du point de vue référencialiste pour arriver à cerner les propriétés fondamentales du mot lui-même. Pour l'atteinte de cet objectif, ils recommandent d'inverser la méthodologie employée habituellement par les dictionnaires :

> *Au lieu de partir d'une description des propriétés de l'objet que le mot catégorise (un type de meuble pour table, un type de végétal pour arbre) pour expliquer comment on peut dériver les autres sens du mot, il faut au contraire tenter de cerner les propriétés du mot lui-même, qui expliquent à la fois qu'il puisse prendre suivant les énoncés ses différents sens, et en même temps qu'il évoque aussi puissamment, à lui tout seul, la classe d'objets à laquelle on l'associe spontanément* (1997 : 51-52).

Polysémie

La sémantique grammaticale, cadre théorique élaboré par Jacques Ouellet, propose une distinction claire entre ces deux aspects sémantiques du mot qui permet d'en expliquer la polysémie. Un mot comporte nécessairement dans l'usage une valeur propre qui permet des références à des faits d'expérience différents en discours selon le contexte où il est employé : il a donc une valeur inhérente en tant que signifié et une valeur référentielle qui tient de l'usage qu'on en fait en syntaxe, en particulier de son association aux autres concepts. Ce cadre théorique présente donc l'avantage de fournir une définition positive des concepts en caractérisant les propriétés intrinsèques qui permettent de les opposer les uns aux autres dans le système de la langue. La valeur intrinsèque d'un concept est stable parce qu'elle découle des composantes et des relations qui sont immanentes à sa structuration[1]. À l'opposé, la valeur référentielle est variable parce qu'elle dépend du contexte auquel est intégrée cette valeur conceptuelle.

[1] Voir *Valeur lexicale et valeur grammaticale* dans le présent ouvrage collectif.

Le phénomène sémantique qu'on désigne communément par le terme de *polysémie* est une propriété des concepts qui se définit justement par le rapport qui s'institue entre sa valeur intrinsèque et sa valeur référentielle dans l'usage. Le rapport qui s'institue entre ces deux valeurs implique donc que la première est nécessairement un facteur de la seconde. Ce rôle que remplit la valeur intrinsèque au sein de cette relation s'explique par le haut degré d'abstraction qui la rend adaptable à différentes références, qui présentent toujours une valeur plus concrète que celle qu'elle évoque.

Pour un même concept, cette référence varie en fonction de la structuration logique du discours où il intervient. Par exemple, la valeur conceptuelle de *arbre*, qui pourrait correspondre abstraitement à l'idée d'un « schème à bifurcations[2] » (DHLF), prend différentes valeurs référentielles selon le contexte d'emploi :

arbre fruitier ;
arbre de transmission ;
arbre de porphyre ;
arbre de la connaissance du Bien et du Mal ; etc.

C'est pareillement le cas de *pomme* qui évoque un objet sphéroïde ressemblant au fruit qui est ainsi désigné[3] :

pomme de terre ;
pomme d'Adam ;
pomme de route ;
pomme McIntosh ; etc.

Ces séquences manifestent chaque fois une exploitation différente d'un même concept – soit « arbre » ou « pomme » selon le cas – en vue de satisfaire aux besoins de la référence.

[2] Compte tenu du haut degré d'abstraction de la valeur conceptuelle, cette définition n'est seulement qu'une approximation. Celle-ci n'est en fait qu'une définition d'ordre référentiel qui se veut la plus représentative de la valeur conceptuelle.
[3] *Idem* à note 2.

Par surcroît, la valeur référentielle de chacune de ces séquences peut même changer en proportion du nombre d'individus qu'elles peuvent caractériser, comme c'est par exemple le cas des trois arbres suivants :

Un signifié doit donc nécessairement comporter une valeur stable pour conserver son identité dans chacune de ses occurrences d'emploi, et c'est pourquoi il est marqué par le même signe. En effet, les mots *arbre* et *pomme*, comme tout autre mot, manifestent une valeur référentielle qui peut varier autant de fois que le permet le contexte. La polysémie implique donc une valeur stable dont le degré d'abstraction permet de produire un nombre variable de valeurs référentielles dans autant de contextes.

La synonymie

Par ailleurs, le fait que des signifiés différents puissent opérer une même référence indique bien qu'ils manifestent ces deux types de valeur sémantique. Par exemple, une phrase de forme passive et une autre de forme active peuvent, malgré leurs différences de composition, renvoyer à un même référent :

Pierre lance la balle.
La balle est lancée par Pierre.

Des signifiés différents peuvent donc renvoyer à des référents identiques mais conceptualisés de manière différente. En effet, la première phrase présente l'événement évoqué par *lance* du point de vue de son lieu d'origine représenté par *Pierre*, tandis que la seconde met l'accent sur le lieu d'échéance de l'événement représenté par le participe *lancée*.

Le cas de référence commune probablement le plus connu est celui de la synonymie dont font état les dictionnaires conçus à cette fin :

> *orgueil, vanité, prétention* (*Petit Robert*, 1996 : *orgueil*) ;
> *sensuel, lascif, voluptueux* (*Idem* : *sensuel*) ;
> *énorme, gigantesque, immense* (*Idem* : *énorme*).

La synonymie se caractérise par la possibilité pour des concepts différents d'opérer une référence similaire ou identique dans l'usage.

La synonymie sert parfois de critère pour assimiler des séquences complexes à des mots en tant qu'expressions figées[4]. C'est par exemple le cas de *pomme de terre* en tant que synonyme de *patate*, de même que celui de *tout à coup* en regard de *soudainement*. Or, on ne peut prétendre que la séquence *pomme de terre* ne constitue qu'un seul mot en prenant pour critère sa synonymie avec *patate* parce que ce type de critère d'ordre référentiel ne rend pas compte de ce qui distingue ces deux unités. Ce critère ne tient pas de la valeur intrinsèque du signifié, mais du contexte d'emploi. Ces deux unités fournissent au locuteur la possibilité de conceptualiser différemment un même référent. En effet, *patate* permet d'envisager le légume en question, tandis que *pomme de terre* permet de l'envisager en tant qu'« aliment comestible végétal dont la forme rappelle celle du fruit et qui vient de la terre[5] ».

De plus, le fait qu'un même mot puisse renvoyer à des référents différents selon l'usage qu'on en fait confirme le fait que le mot comporte une valeur inhérente distincte de sa valeur référentielle.

Le rapport sémantique qui s'institue entre l'hyperonyme et l'hyponyme peut dans certains emplois correspondre à une relation de synonymie. C'est notamment le cas du nom *conifère* auquel on peut substituer l'hyperonyme *arbre* ou l'hyponyme *sapin* pour opérer la même référence :

---

[4] Voir Parent (1999).
[5] Voir infra : 7, note de bas de page 2.s

Un *conifère* pousse devant la maison ;
Une *arbre* pousse devant la maison ;
Un *sapin* pousse devant la maison.

L'équivalence sémantique entre le terme et sa définition, plus précise que la précédente, constitue également un rapport de synonymie :

homme : *animal raisonnable*.

En fonction des besoins de son expression, un locuteur choisit les valeurs conceptuelles qui satisfont à la référence qu'il veut opérer. Ainsi, les ressources conceptuelles dont dispose la langue fournissent-elles au locuteur le moyen de s'exprimer différemment au sujet d'un même référent en mettant à l'avant-plan l'un ou l'autre de ses aspects. Il peut par exemple envisager un même arbre d'un point de vue général (conifère) ou du point de vue de sa spécificité (sapin).

Conclusion

Le mot est donc un signe qui codifie un concept, lequel dans l'usage en discours renvoie à un référent particulier.

De son côté, le mot en tant que signe n'entretient pas de relation organique avec le signifié qu'il codifie, contrairement à ce que laisse entendre Saussure lorsqu'il avance que le signe unit un concept et une image acoustique (1975 : 98) : cette « union » n'en fait pas une entité organique, puisque la fonction du signe à l'égard du signifié est d'ordre strictement symbolique. Autrement dit, il n'existe dans le cadre du mot ni rapport causal ni rapport analogique entre le signe et le signifié.

Le concept, pour sa part, manifeste une valeur inhérente qui conditionne sa valeur référentielle. Cette valeur conceptuelle est stable parce qu'elle dépend de la valeur des composantes qui le structurent. Elle requiert donc une définition opératoire du concept, ce que propose la sémantique grammaticale qui avance

que le concept résulte de la détermination d'une composante grammaticale par une composante lexicale[6]. La distinction de la valeur conceptuelle de la valeur référentielle est nécessaire à la réalisation d'une analyse adéquate dans le domaine de la sémantique. Cette distinction permet notamment de mettre à jour la relation de cause à effet qui s'institue entre la valeur conceptuelle et la valeur référentielle, ce qui permet de jeter un regard neuf sur le phénomène de la polysémie dont on ne parvient généralement pas à expliquer autrement que d'un point de vue centré sur la référence. Cette distinction permet également de jeter un regard neuf sur la synonymie qui se définit en tant que phénomène de concordance référentiel qui dépend du contexte d'emploi. En outre, elle conditionne la formulation d'une définition des catégories de mots qui convienne à l'ensemble des membres qu'elles incluent. Elle évite donc les problèmes engendrés par une analyse centrée sur la référence ou sur la syntaxe, et qui mènent à des catégories non valides.

Finalement, le référent permet d'évoquer figurativement aussi bien une chose qui appartient au monde objectif (table, chaise, etc.) qu'une chose qui appartient au monde imaginaire (licorne, griffon, etc.). C'est pourquoi la cohérence d'une analyse sémantique requiert pareillement de bien distinguer la valeur référentielle d'ordre figuratif de l'objet physique ou imaginaire qu'elle représente.

Bibliographie

DE MAURO, T. (1969), *Une introduction à la sémantique*, Paris, Payot.

DE SAUSSURE, F. (1975), *Cours de linguistique générale*, Paris, Payot.

[6] Voir « Valeur lexicale et valeur grammaticale » dans le présent ouvrage collectif. Voir également Ouellet (1990, 1985 et 1979).

[DHLF] : ROBERT, P. (1998), *Le Robert, Dictionnaire historique de la langue française*, sous la direction d'Alain Rey, Paris, Dictionnaire Le Robert.

OUELLET, J. (2000), *Précis de sémantique grammaticale*, Note de cours, Québec, Université Laval, Faculté des Lettres, Département de langues, linguistique, et traduction.

OUELLET, J. (1990), « Systématique de représentation spatiale », *Langues et linguistique*, 16, Québec, Université Laval, p. 149-175.

OUELLET, J. (1985), « Sémantique grammaticale du nom et du pronom », *Langues et linguistique*, 11, Québec, Université Laval, p. 191-239.

OUELLET, J. (1979), « Constitution syntagmatique des unités nominales », *Langues et linguistique*, 4-5, Québec, Université Laval, p. 115-140.

PARENT, F. (2000), « Sémantique grammaticale du mot tout en français contemporain ». Mémoire, Québec, Université Laval.

PARENT F. (2000), « Valeur lexicale et valeur grammaticale », *Actes des XIV^e Journées de Linguistique*, Québec, Ciral.

PARENT F. (1999), « De l'inexistence de locutions », *Actes des XIII^e Journées de Linguistique*, Québec, Ciral, p. 165-184.

SCHAFF, A. (1969), *Langage et connaissance : suivi de six Essais sur la philosophie du langage*, Paris, Anthropos.

VAN VLASSELAER, J.-J. (1973), « Délimitation et définitions du mot à travers les différentes conceptions de sciences linguistiques après F. de Saussure ». Mémoire de maîtrise, Ottawa, Université d'Ottawa.

VICTORRI, B., et C. FUCHS (1996), *La polysémie, construction dynamique du sens*, Paris, Hermès.

WHORF, B. L. (1971), *Linguistique et anthropologie*, Paris, Denoël.

Valeur lexicale et valeur grammaticale

François Parent
Université Laval

Introduction

En linguistique, on distingue généralement la valeur lexicale et la valeur grammaticale des mots, au moins dans les cas où les composantes grammaticales sont marquées par la sémiologie. Cependant, dans la pratique de l'analyse sémantique, on n'identifie pas clairement et on ne décrit pas explicitement ces deux composantes. Ce défaut de méthode entraîne des lacunes considérables dans le classement des mots. Souvent, ce classement se fonde sur la valeur lexicale du mot, celle qui est marquée, parce qu'on ne dispose pas d'une méthode qui permette de reconnaître clairement la composante grammaticale, celle qui n'est souvent pas marquée par le signe. Or, une même composante lexicale peut entrer dans la composition de concepts de nature différente : *le **deux** de cœur* ; ***deux** hommes*. Elle ne peut donc constituer ni un critère de catégorisation valable, ni un critère de distinction des mots appartenant à des catégories différentes. Ceux-ci peuvent avoir la même composante lexicale marquée par le même signe – *deux* évoque l'idée de « deux » dans les trois cas – mais ne pas avoir la même structuration grammaticale, ce qui entraîne des problèmes de classement impossible à résoudre si l'on néglige la valeur sémantique caractéristique de chacune des diverses catégories en cause. Il faut donc pour caractériser adéquatement les différentes parties du discours trouver le moyen d'identifier et de décrire la valeur catégorielle qui caractérise les différents types de mots même dans les cas où ils comportent la même composante lexicale.

La présente communication vise donc à illustrer comment, dans le cadre théorique de la sémantique grammaticale développé par

Jacques Ouellet, on arrive à identifier et à décrire le signifié grammatical, ainsi que la relation qu'il entretient avec le signifié lexical pour structurer le mot en vue d'établir un classement des parties du discours fondé sur des critères explicites adéquats. Cette méthode d'analyse sémantique met en cause l'extension – c'est-à-dire la représentation morphologique de l'espace (voir Ouellet, 1990) – en tant que facteur de structuration fondamental des concepts. L'analyse de ce facteur morphologique et de son rôle dans la structuration des concepts permet de définir les structures conceptuelles qui conditionnent la structuration grammaticale et logique de la phrase. Cette analyse s'effectue sur deux plans : celui de la paradigmatique où elle porte sur les composantes grammaticales des concepts, et celui de la syntagmatique où elle met en cause les relations entre les diverses composantes qui entrent dans leur structuration.

Le concept

Comme il appartient toujours à une catégorie, mais qu'il se distingue toujours de toutes les autres unités de sa catégorie, le concept marqué par le mot implique nécessairement la combinaison d'un signifié lexical qui lui donne sa valeur spécifique et d'un signifié grammatical qui lui donne sa valeur catégorielle. La composante lexicale est formée d'unités concrètes comme les lexèmes (*pomme*, *garde-fou*) ou d'unités plus abstraites comme les morphèmes lexicaux (pomm*ier*, anthropo+log*ie*, *re*partir), et la composante grammaticale ne comporte que des unités très abstraites dites morphèmes grammaticaux comme le genre, le nombre et l'aspect, par exemple : (ch*iens* / ch*ien* / chien*ne* / chien*nes* ; march*ais*, march*a* ; etc.). Cependant, faute d'être marquées par le signe, ces composantes grammaticales n'ont pas toutes été identifiées par l'analyse traditionnelle et l'analyse linguistique moderne. C'est pourquoi encore aujourd'hui on tend à recourir à la syntaxe pour définir certains types de mots plutôt que de se fonder sur leur constitution propre. On tente donc de les

définir non pas sur la base de ce qu'ils sont intrinsèquement, mais sur la base de leur rôle dans la phrase. En conséquence, on a tendance à confondre nature et fonction.

La composante lexicale

Le signifié lexical manifeste la valeur distinctive qui oppose un concept aux autres concepts de sa catégorie. Le lexème y représente une valeur particulière et relativement concrète par rapport aux morphèmes qui ont une valeur générale et plus abstraite.

Le lexème

La langue fournit une quantité non limitée de lexèmes qui permettent la structuration d'autant de concepts de valeur distincte. Par exemple, chacun des lexèmes qui entrent dans la composition des noms caractérisant le domaine des couleurs : *rouge*, *jaune*, *vert*, *bleu*, *violet*, etc., évoque une notion particulière qui se définit par opposition aux autres. Il s'agit de valeurs mutuellement exclusives : *rouge* n'est ni *jaune*, ni *vert*, ni *bleu*, ni *violet*, mais *rouge*. De même que *jaune* n'est ni *rouge*, ni *vert*, ni *bleu*, ni *violet*, mais *jaune*. Cette identification par opposition spécifique s'étend à tous les lexèmes disponibles dans le domaine lexical pris dans son ensemble. En effet, *rouge* n'est ni *grand*, ni *nombre*, ni *table*, ni *pêche*, mais *rouge*.

Les lexèmes sont des composantes substituables en série ouverte pour une même fonction dans le cadre d'un concept, et qui sont toujours marquées par un signe distinctif : con*cern*ant, con*cert*ant, con*firm*ant, con*clu*ant, con*gru*ant, con*jur*ant, etc. De même, *table* et *pêche* n'ont pas la même valeur conceptuelle et ne renvoient pas à la même réalité dans :

a) La *table* est tombée.
b) La *pêche* est tombée.

à cause de la valeur spécifique de leur lexème. En effet, le concept « table », qui renvoie dans cette phrase au référent qu'on décrit comme un « meuble fait d'un plateau horizontal posé sur un ou plusieurs pieds » (*Larousse*, 1987) diffère du concept « pêche » qui renvoie en ce cas au « fruit comestible à noyau du pêcher » (*idem*).

Le signifié lexical peut être constitué d'un seul lexème, ce qui est par exemple le cas de *table*, ou de plus d'un lexème dans des mots composés comme *garde-fou*, *porte-avion* ou *casse-noisette*.

Le morphème lexical

Comme les morphèmes grammaticaux, les morphèmes lexicaux relèvent de paradigmes qui associent un nombre limité de morphèmes commutant en série fermée. Ce type de morphèmes représente des notions générales plus abstraites que les lexèmes. En conséquence, un signifié lexical qui comprend uniquement des morphèmes est de valeur plus abstraite que celui qui comporte un lexème. C'est particulièrement le cas de certaines prépositions : *de*, *dans*, *à*, *en* ; et de certains pronoms : *je*, *tu*, *le*, *mon*, dont la valeur est plus abstraite que celle de noms comme *pomme*, *bleu*, *pierre*, *arbre*, etc. Le morphème lexical correspond sémiologiquement à un préfixe : *antirouille*, *antidote*, ou à un suffixe : *tablette*, *noisette*, selon qu'il précède ou suit le lexème dans la formation d'un concept. Le signifié lexical du concept peut donc comporter uniquement un morphème lexical, ce qui est le cas de l'article, par exemple, qui oppose sous ce rapport *un* et *le*. Il peut également associer un lexème et un morphème lexical (souligné) dans des structures comme *tabl-ette*, *anti-rouille* ou *nois-ette*, ou comporter plus d'un morphème lexical dans des séquences comme : *agré-able-ment*.

Le signifié lexical qui distingue un concept des autres unités de sa catégorie peut entrer dans la structuration des concepts de nature différente puisque sa valeur est indifférente à la catégorisation grammaticale. Ainsi, le lexème *tout-* peut entrer dans la composition du nom *tout* (*le* tout *est plus grand que la partie*), dans celle du pronom *tout* (tous *les hommes*), ou dans celle de l'adverbe (*elles sont* tout *agitées*). Ce qui est le cas des morphèmes lexicaux également : un même morphème lexical intervient aussi bien dans la structuration d'un verbe, que dans celle d'un nom, ou d'un participe : *dé*tourner, *dé*tour, *dé*tournement, *dé*tourné. Ce n'est donc pas la composante lexicale, mais bien sa composante grammaticale et sa structuration qui déterminent la nature d'un concept.

La composante grammaticale

Le signifié grammatical manifeste les notions qui caractérisent un ensemble de concepts et en fondent la catégorisation. Les paradigmes grammaticaux sont donc des systèmes de langue qui définissent les notions générales communes aux différents concepts d'une même catégorie. Par exemple, l'analyse de la morphologie du nom révèle que ce type de concepts exploite dans tous les cas un morphème d'extension, un morphème de genre et un morphème de nombre[1].

Les morphèmes grammaticaux qui appartiennent à un même paradigme présentent différentes versions d'une même notion et se substituent les uns aux autres dans la formation d'un concept. Ces morphèmes qui constituent une série fermée s'opposent les uns aux autres et s'excluent mutuellement pour une même fonction. Ce qui fait qu'un même concept ne peut comporter simultanément les morphèmes de nombre singulier et pluriel ou ceux de genre féminin et masculin. Ces morphèmes grammaticaux se distinguent des composantes lexicales par le fait qu'il est obligatoire d'employer

[1] Voir Ouellet (1985), Audet (1994 : 42) et Parent (2000 : § 2.8.1).

l'une des notions d'un paradigme pour la formation d'un type de concept. Ainsi, par exemple, on ne trouve jamais de nom en français qui ne comporte pas de morphème de genre, de morphème de nombre ou de morphème d'extension.

La paradigmatique

La paradigmatique met en cause les systèmes de langue qui permettent de produire les notions qui composent les différentes catégories de concepts. Étant donné que les morphèmes d'un même paradigme appartiennent tous à un même système, l'analyse paradigmatique doit déterminer la valeur sémantique de chacun en se fondant sur les oppositions qui se manifestent dans le cadre de ce système. Cette analyse permet donc de décrire la valeur des morphèmes grammaticaux qui entrent dans la structuration d'un concept et la structuration des paradigmes qui les définissent.

L'analyse de la composante grammaticale des concepts contribue à identifier la catégorie à laquelle ils appartiennent sur la base de critères pertinents. Ainsi, par exemple, la présence d'un genre et d'un nombre exclut un concept de la catégorie du verbe, de celle de l'adverbe et de celle de la préposition, car la structuration de ces types de concept n'exploite pas cette morphologie. Toutefois, un tel concept peut appartenir aussi bien à la classe des noms qu'à celle des pronoms qui comportent cette morphologie.

L'extension

On admet généralement, comme l'enseigne la logique, que le nom a une extension, et les grammairiens font état de cette propriété : « son extension, ou l'étendue de la signification du nom, est le nombre plus ou moins grand des êtres auxquels l'idée peut s'appliquer » (Grevisse, 1980 : § 383). Par exemple, le nom *animal* est de plus large extension que *chat*. En effet, tous les chats sont des

animaux, mais touss les animaux ne sont pas des chats. En d'autres termes, l'ensemble des chats constitue un sous-ensemble de l'ensemble des animaux. Un nom peut donc avoir une extension plus ou moins considérable selon sa valeur lexicale, c'est-à-dire désigner un ensemble plus ou moins considérable d'individus.

Cependant, les grammairiens ne reconnaissent pas l'extension comme un facteur sémantique appartenant au plan de la morphologie grammaticale, soit parce que ce facteur n'est pas marqué par le signe, soit parce que par principe on exclut de la grammaire tout facteur d'ordre sémantique. Mais on postule en sémantique grammaticale que la grammaire tient fondamentalement de la structuration sémantique d'une langue et on pose qu'un facteur sémantique qui est commun à toutes les unités d'une catégorie de concepts appartient à sa morphologie grammaticale. Dans le cas du nom, ce critère permet d'identifier non seulement le genre et le nombre, mais aussi l'extension : tous les noms comportent sans exception une extension, c'est-à-dire une représentation abstraite d'espace qui joue un rôle fondamental dans la structuration du concept. Il s'agit de la représentation abstraite de l'ensemble des individus auxquels peut s'appliquer une détermination lexicale donnée :

> *Cette conceptualisation implique la mise en rapport d'un univers idéel contenant quantitativement variable et d'un contenu individualisé qui tient lieu d'élément de cet ensemble ; il s'agit d'une part de la représentation abstraite d'un espace d'étendue plus ou moins considérable qui fait office de lieu d'incidence, et, d'autre part, de l'évocation également abstraite de l'individu qui tient lieu de contenu incident à ce lieu d'existence* (Ouellet, 1985 : 201).

L'espace contenant ainsi évoqué est quantitativement variable : il a alternativement dans l'usage la dimension d'un individu ou celle

d'un groupe plus ou moins considérable. Le nom *chat* par exemple, peut renvoyer à un ensemble qui inclut un seul individu (en a), un groupe plus ou moins considérable d'individus (en b) ou encore, à tout individu possédant les mêmes caractéristiques (en c) :

a) Le *chat* de ma sœur est blanc et noir.
b) Elle a eu plusieurs *chats* dans sa vie.
c) Le *chat* est un animal.

L'extension se définit par opposition à la compréhension qui est « le nombre plus ou moins grand des notes ou éléments que comprend cette idée » (Grevisse, 1980 : § 384), ce qui correspond aux propriétés spécifiques de ce qui appartient à l'ensemble évoqué par le nom. La compréhension correspond donc à la valeur lexicale du concept. En conséquence, *chat* est un nom de moins grande extension et de plus grande compréhension que *animal* : il met en cause un moins grand nombre d'individus que ce dernier.

Cette représentation d'un ensemble, peu importe son étendue, est une caractéristique commune à tous les concepts qui appartiennent à la catégorie du nom. C'est pourquoi l'extension est une propriété grammaticale du nom, non une propriété lexicale comme la compréhension qui représente ce qui est spécifique à un concept. Cette valeur lexicale permet d'opposer les noms les uns aux autres parce qu'elle confère à l'ensemble que chacun représente un contenu spécifique différent. D'un nom à un autre, cet ensemble inclut donc des éléments de nature différente, ce qu'illustrent des concepts comme « chat », « chien », « arbre », « sapin », « droite », « gauche », etc.

Par contre, certains concepts ne caractérisent pas les éléments de l'ensemble qu'ils représentent, mais comportent une détermination lexicale qui le caractérise en tant que contenant, ce qui leur confère leur nature de pronom[2]. Dans ce cas, le signifié lexical détermine

[2] Elle fonde également celle du verbe. Mais vu la complexité de ce concept, les

une propriété générique d'un ensemble, non pas une propriété spécifique de ses éléments. Le contenu de cet ensemble demeure donc indéterminé dans le cadre du concept. Le pronom s'oppose donc au nom par sa structuration syntagmatique et ce, même lorsque ces deux types de concepts comportent les mêmes composantes.

e) À moi, Comte, *deux* mots (Corneille dans Grevisse, 1980 : § 882).
f) Les *dix* commandements.

Ainsi, les pronoms *deux* et *dix* confèrent à l'ensemble que chacun représente une valeur quantitative en tant que contenant sans spécifier la qualité des éléments de cet ensemble. C'est pourquoi cette spécification exige la mise en rapport du pronom numéral avec des noms comme *mots* et *commandements*, qui évoquent la qualité spécifique de ce qui appartient à l'ensemble que chacun représente. Cependant, les noms *deux* et *dix* dans la séquence suivante :

g) Le *deux* de cœur et le *dix* de pique.

n'évoquent pas la quantité mais la qualité spécifique du contenu de l'ensemble que chacun représente. Cette valeur qualitative est par ailleurs comparable à celle de noms, comme *valet*, *dame*, *roi* et *as* qui jouent un rôle analogue à celui de *deux* et *dix* dans la séquence précédente :

h) Le *valet* de cœur et la *dame* de pique.
i) Le *roi* de cœur et l'*as* de pique.

Pour quantifier les éléments contenus de l'ensemble que le nom numéral représente, on l'associe comme tout autre nom à un pronom quantificateur en syntaxe. C'est le cas du nom numéral

---

limites de cet article ne permettent pas d'en discuter. Pour plus de précision, voir Ouellet (1987 ; 1988).

deux qui, comme dans le cas d'*épingles*, est associé au pronom *quatre* dans :

j) Être tiré à *quatre épingles* (*Petit Robert*, 1996 : s.v. *quatre*).
k) Il y a *quatre deux* dans un jeu de cartes : le deux de cœur, le deux de pique, le deux de carreau et le deux de trèfle.

La valeur qualitative des noms numéraux les oppose donc aux pronoms numéraux de valeur quantitative bien qu'ils comportent tous les deux une même composante lexicale : ce qui démontre que la fonction de la composante lexicale contribue considérablement à la valeur sémantique du concept.

Enfin, la détermination lexicale peut ne déterminer ni la qualité spécifique des éléments d'un ensemble ni une propriété de cet ensemble en tant que contenant, mais porter sur le rapport d'appartenance qui le définit. Cette détermination a pour effet d'évoquer une modalité ou un mode d'appartenance à cet ensemble, ce qui est typique de l'adverbe et de la préposition, où la composante lexicale en détermine en quelque sorte le mode d'existence. Ainsi, par exemple, l'adverbe *fort* qui détermine une modalité de *intéressant* dans : « c'est *fort* intéressant » s'oppose à l'adjectif nominal *fort* qui détermine une qualité spécifique de *homme* dans « un homme *fort* ».

Il apparaît ainsi que la catégorisation grammaticale, si elle ne tient pas de la valeur lexicale du concept, met clairement en cause la valeur sémantique qui est liée à la fonction de cette composante dans la structuration du concept, une fonction qui est nettement différente et qui produit une valeur conceptuelle différente selon qu'on se trouve en présence d'un nom, d'un pronom ou d'un adverbe ainsi que l'illustrent à l'évidence les concepts analysés.

Conclusion

Tout concept se structure par l'association d'une composante lexicale de valeur particulière et d'une composante grammaticale de valeur catégorielle. Dans la mesure où tout concept implique la représentation d'un ensemble qui évoque une portion d'espace plus ou moins considérable, tout concept comporte nécessairement cette composante grammaticale, c'est-à-dire la représentation d'une relation d'appartenance entre un contenant et un contenu. Un concept dont la détermination lexicale porte sur le contenu, c'est-à-dire les éléments de l'ensemble constituent un nom. Si cette détermination porte sur l'ensemble en tant que contenant, il s'agit d'un pronom. Si cette détermination porte sur la relation d'appartenance qui définit cet ensemble, le concept est selon le cas un adverbe. La même composante lexicale peut donc évoquer une qualité, une quantité ou une modalité selon sa fonction dans la formation du concept, ce qui fournit en chaque cas un critère sémantique explicite pour fonder le classement de ces concepts.

Bibliographie

AUDET, C.-H. (1994), *Morphologie et syntaxe du français. Précis théorique et méthodologique de sémantique grammaticale*, Sainte-Foy, Le Griffon d'argile.

GREVISSE, M. (1980), *Le bon usage. Cours de grammaire française et de langage français*, 11^{e} éd., Gembloux et Paris, Duculot.

OUELLET, J. (1990), « Systématique de représentation spatiale », *Langues et linguistique*, 16, Québec, Département de langues et linguistique, Faculté des Lettres, Université Laval, p. 149-175.

OUELLET, J. (1988), « Sémantique grammaticale du verbe II », *Langues et linguistiques*, 14, Québec, Département de langues et linguistique, Faculté des Lettres, Université Laval, p. 199-249.

OUELLET, J. (1987), « Sémantique grammaticale du verbe I », *Langues et linguistiques*, 13, Québec, Département de langues et linguistique, Faculté des Lettres, Université Laval, p. 183-230.

OUELLET, J. (1985), « Sémantique grammaticale du nom et du pronom », *Langues et linguistique*, 11, Québec, Département de langues et linguistique, Faculté des Lettres, Université Laval, p. 191-239.

OUELLET, J. (1979), « Constitution syntagmatique des unités nominales », *Langues et linguistique*, 4-5, Québec, Département de langues et linguistique, Faculté des Lettres, Université Laval, p. 115-140.

PARENT, F. (2000), *Sémantique grammaticale du mot* tout *en français contemporain*. Mémoire, Québec, Université Laval.

PARENT, F. (1999), « De l'inexistence de locutions », Actes des XIII[e] Journées de Linguistique, Québec, Ciral, p. 165-184.

Petit Larousse Illustré (1986), Paris, Larousse.

ROBERT, P. (1996), *Nouveau Petit Robert. Dictionnaire alphabétique et analogique de la langue française*, rédaction dirigée par Alain Rey et Josette Rey-Debove, Paris, Dictionnaires le Robert.

Vues d'aujourd'hui sur le patrimoine

Natacha Auclair, Pascale Bélanger,
Stéphanie Campeau, Geneviève Fradette,
Michel Jr Laflamme, Marie-France LeBlanc,
Patricia Pelletier, Jean-François Smith,
Marie-Hélène Tremblay
Université Laval

À l'automne 1999, madame Diane Vincent, professeure d'analyse du discours à l'Université Laval, nous a proposé un travail de recherche sur la perception du patrimoine dont les résultats seraient utilisés à l'occasion d'une exposition à Deschambault. Nous avons travaillé avec un questionnaire déjà établi et validé par madame Vincent. Il comportait trois questions : « selon moi, le patrimoine c'est ... », « selon moi, le patrimoine ça sert à ... » et « ce qui représente le mieux le patrimoine dans ma vie, c'est ... ».

Chacune des trois questions portait sur un aspect différent de la notion de patrimoine : le concept, la fonction et les exemples. Ces trois aspects sont traités successivement dans l'une ou l'autre des sections du texte avec, en dernier lieu, une réflexion éclairée sur la définition du patrimoine.

La cueillette des données a été faite selon la méthode boule-de-neige. Nous avons dépouillé 378 questionnaires : 135 complétés par des hommes et 238 par des femmes. Le tableau suivant illustre la distribution par groupe d'âge :

Tableau 1 : Distribution par groupe d'âge

Groupe d'âge	Nombre de répondants
10-17 ans	59
18-25 ans	123
26-35 ans	47
36-55 ans	94
56-85 ans	52

Précisons que même si la répartition est inégale, tous les groupes sont bien représentés.

Pour compiler les données, nous avons fait une pré-analyse à l'aide de 25 questionnaires, ce qui nous a permis de créer des catégories de classement. Puis, nous avons consulté le découpage gouvernemental de la notion de patrimoine tant au fédéral qu'au provincial, pour constater que leurs catégories étaient insatisfaisantes. Nous avons donc conservé nos propres catégories et les avons validées avec 25 autres questionnaires. Quelques ajustements ont été faits en cours de route, afin de procéder à la codification de toutes les réponses.

Problèmes de définition

La conceptualisation dans un but définitionnel est un exercice ardu, et ce, même pour des lexicographes expérimentés. De plus, il est beaucoup plus difficile de définir des idées, des concepts, qui n'ont pas de signifié matériel, que des objets : définir le mot *œil* n'est pas du même ordre que de définir *voir*, par exemple. Ayons donc une bonne pensée pour les répondants, *patrimoine* n'étant pas de tout repos à définir.

Sans entrer trop profondément dans la lexicologie, voyons d'abord ce qu'est une définition. Dans un sens restreint, la définition porte sur le concept. Comme l'explique Jean-Claude Boulanger : « la définition est une équation sémique qui se représente sous la forme d'une paraphrase synonymique de l'entrée » (1997).

> *La définition est un énoncé qui décrit prioritairement l'ensemble des traits sémantiques appartenant au mot et qui renseigne sur l'une ou l'autre des acceptions (sens) qu'il peut avoir.* [...] *Accessoirement, la définition incorporera des informations encyclopédiques qui traitent du référent (ndr : lire ici : l'utilité et les exemples)* (1997).

En fait, la définition ne porte que sur la relation signe-objet ; elle décrit le sens du mot et non pas l'objet désigné par le mot. On reste donc dans un métadiscours qui n'a aucune relation concrète avec le monde physique désigné par le langage. Lorsque l'on fait intervenir le « à-quoi-ça-sert », on passe alors dans un autre contexte, celui de la relation de l'objet désigné avec le monde auquel il appartient. On quitte alors le métadiscours pour entrer dans la définition encyclopédique. Cela est encore plus vrai lorsque l'on donne des exemples : on réfère directement au signifié, à « l'objet dans le monde ». À partir de ces remarques, on peut présumer que chaque question du questionnaire servira, d'une certaine façon, à définir le patrimoine.

Pour la première question, qui était, rappelons-le, « Selon moi, le patrimoine c'est ... », la majorité des réponses que nous avons reçues ne comprenaient pas l'aspect conceptuel. Nous avons eu énormément de listes d'exemples, beaucoup de descriptions des fonctions du patrimoine, mais peu de définitions proprement dites. Précisons que les exemples obtenus demeuraient des génériques, les spécifiques se retrouvant à la troisième question. Par exemple, bien des gens ont parlé de l'architecture à la première question et ont donné une église ou un bâtiment historique en exemple à la troisième question. Le tableau suivant illustre chacun des constituants des définitions.

Tableau 2 : Les constituants des définitions

Constituant	Exemple
Concept	L'ensemble des coutumes, pratiques, institutions propres à une culture et à une tradition qui a perduré au fil de l'histoire. (Q.73)
Fonction	Laisser à notre relève une partie de notre histoire, d'où l'on vient. Faire valoir nos droits, nos coutumes et d'où elles partent. Faire connaître aux autres notre identité autant culturelle que familiale. Une identification. (Q.63)
Exemple	La langue, les valeurs, notre demeure, notre pays, notre mode de vie, « nos mœurs ». (Q.53)

Après avoir analysé les réponses aux questions sur la fonction du patrimoine et ses exemples, nous reviendrons sur une définition du patrimoine, celle-ci étant créée à l'aide des éléments de réponse.

La fonction du patrimoine

La deuxième question à laquelle les répondants devaient répondre était : « Selon moi, le patrimoine ça sert à ... ».

Cette question devait pouvoir amener les gens à approfondir leur conception du patrimoine en décrivant son utilité. Mentionnons au passage la conséquence que cette formulation entraîne sur la nature des réponses recueillies : ces dernières sont centrées sur un verbe. C'est pourquoi chacune des sept catégories est représentée par un verbe générique :

a) *se souvenir*, plus spécifiquement des choses, des gens ou de leurs racines (ou origines), comme l'attestent respectivement ces trois exemples :
« Se rappeler les générations antérieures ». (Q.296)

« Il [le patrimoine] sert à raconter et décrire les us et coutumes dépendant de l'époque et de l'environnement où les gens évoluaient ». (Q.284)
« Ça sert à se rappeler de nos origines [...] ». (Q.330)

b) *s'identifier*, en tant qu'individu ou collectivité :
« [...] donner une identification, une identité à une personne ou à un groupe. C'est notre richesse individuelle ou collective ». (Q.379)
« [...] fournir et préserver un sentiment d'identité [...] pour une collectivité dans le sens large du terme, ou pour de petits groupes ». (Q.287)

c) *préserver* :
« Conserver notre histoire ». (Q.419)

d) *transmettre* :
« Transmettre les acquis, [le] savoir-faire et la créativité des générations passées et actuelles (d'un peuple) aux générations futures ». (Q.237)

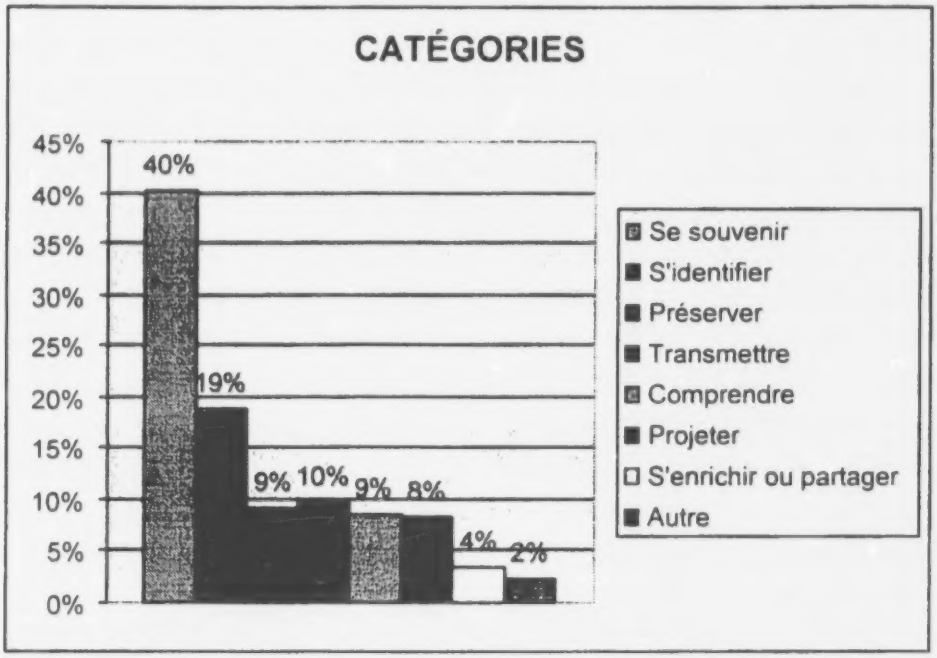
e) *comprendre* :
« Sert à connaître le passé des choses et pourquoi aujourd'hui les choses sont ainsi ». (Q.423)

f) *se projeter* dans l'avenir, dans le sens où le patrimoine sert de leçons pour les générations futures :
« Nous aider à voir où on va ». (Q.363)

g) *s'enrichir* ou *partager* :
« Attirer les touristes ». (Q.11)
« Partager les biens entre deux personnes ». (Q.19)

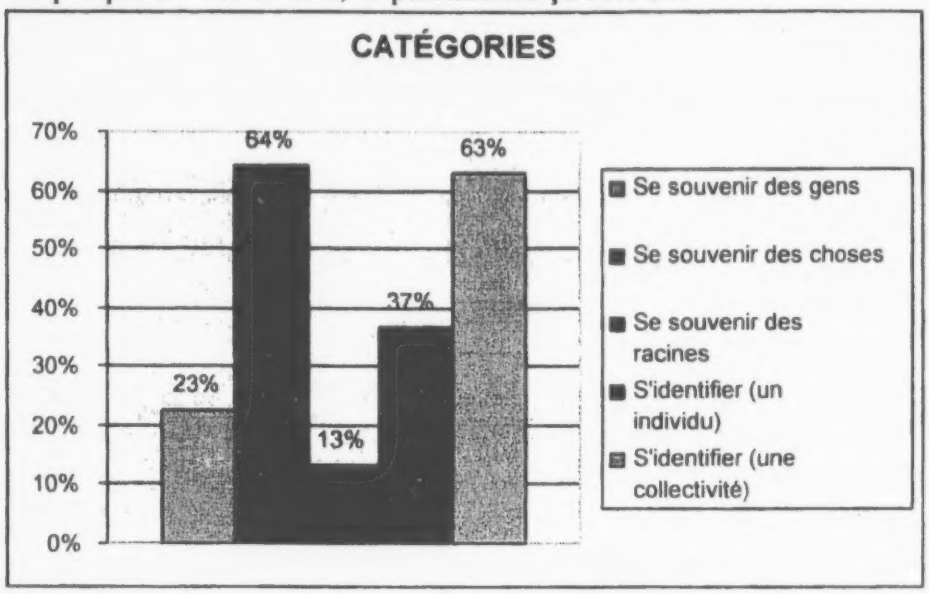
Le graphique 1 illustre la répartition des 548 réponses dans les 7 catégories précédemment exposées. 40,3 % des réponses portent sur le souvenir et 18,8 % sur l'identification. Ces deux catégories représentent à elles seules 59,1 % du total des réponses.

Graphique 1 : Selon moi, le patrimoine ça sert à ...



Le graphique 2 présente un raffinement des deux catégories les plus importantes quantitativement : la catégorie *se souvenir* se subdivise en trois sous-catégories (les gens, les choses et les racines), et la catégorie *s'identifier* se subdivise en deux sous-catégories (une personne, une collectivité). 64,3 % des réponses, une très large majorité, se rapportent au souvenir des choses. C'est le souvenir d'objets, mais aussi – sinon surtout – le souvenir d'actes (« les anciennes habitudes de vie des anciens » (Q.266)) et d'événements passés, « choses » plus immatérielles. *S'identifier*, en tant que collectivité, récolte quant à lui 61,3 % des réponses.

Graphique 2 : Selon moi, le patrimoine ça sert à...



Si nous mettons en relation les réponses obtenues avec le concept de temps, nous constatons que le patrimoine ne fait pas uniquement partie du passé. En effet, *s'identifier*, *comprendre*, *s'enrichir* et *partager* sont quatre types de réponses ancrées dans le présent, alors que *projeter*, *transmettre* et *préserver* s'orientent vers le futur. La catégorie *comprendre* est en réalité à cheval sur le passé et le présent (« le passé peut souvent expliquer le présent » (Q.431)) ; la catégorie *projeter* (« tirer des leçons et construire l'avenir » (Q.222)) établit un pont entre le passé et le futur. Bref, il n'y a que le souvenir (40 % des occurrences) qui soit tourné uniquement vers le passé.

La représentation du patrimoine

Les gens interrogés ont dû répondre, en dernier lieu, à la question suivante : « Qu'est-ce qui représente le mieux le patrimoine dans ma vie ? ». Contrairement aux deux questions précédentes, celle-ci visait des éléments de réponse plus concrets. 906 éléments de réponse ont été comptabilisés et classés dans l'une ou l'autre des catégories énumérées et illustrées dans le tableau 3.

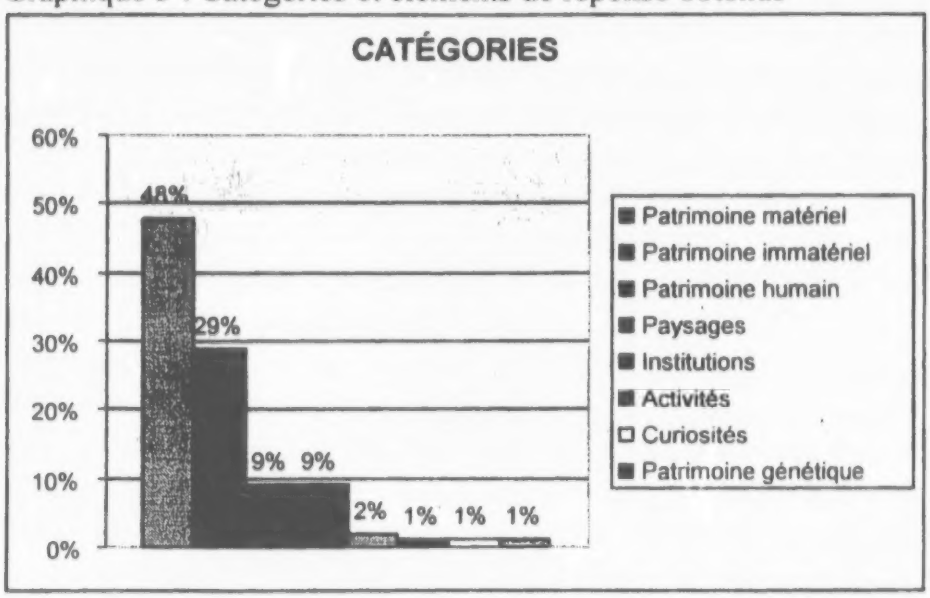
Tableau 3 : Catégories et exemples d'éléments de réponse

1. PATRIMOINE MATÉRIEL
Bâtiments : la basilique Sainte-Anne-de-Beaupré, le Château Frontenac, la vieille Pulperie de Chicoutimi
Biens matériels : une armoire, un document olographe, une bague
Biens emblématiques : une ceinture fléchée, le drapeau du Québec
2. PATRIMOINE IMMATÉRIEL
Culture (sens générique) : la culture
Arts : la littérature, la musique, le cinéma
Valeurs : l'honnêteté, le respect, le sens de la famille
Langue : le français, le québécois
Tradition : les coutumes laissées par nos ancêtres, la bénédiction paternelle
Religion : la religion
Autres : les faits historiques marquants
3. PATRIMOINE HUMAIN
Générique : les ancêtres, les vieux
Spécifique : mes parents, mes enfants, Félix Leclerc, Maurice Richard
4. PAYSAGES
Aménagés : les réserves indiennes, les Plaines d'Abraham
Naturels : le fleuve St-Laurent, la forêt
5. INSTITUTIONS
Les musées, les écoles
6. ACTIVITÉS
La cabane à sucre, le Carnaval de Québec
7. CURIOSITÉS
Les Chinois qui se promènent sur les Plaines d'Abraham
8. PATRIMOINE GÉNÉTIQUE
Mes yeux verts, mes cheveux bruns

Les réponses regroupées dans la catégorie *patrimoine matériel* sont beaucoup plus spécifiques que celles relevant du *patrimoine immatériel*. Ainsi, les répondants ont généralement nommé l'église de leur paroisse, l'école de leur quartier ou encore des biens matériels précis, mais ont rarement mentionné des oeuvres, des valeurs ou des traditions spécifiques. À titre

d'exemple, seulement deux répondants ont nommé des faits langagiers spécifiques : « ch'comprenerrions pas » et « l'emprunt au domaine maritime *embarquer* dans l'autobus ».

Graphique 3 : Catégories et éléments de réponse obtenus

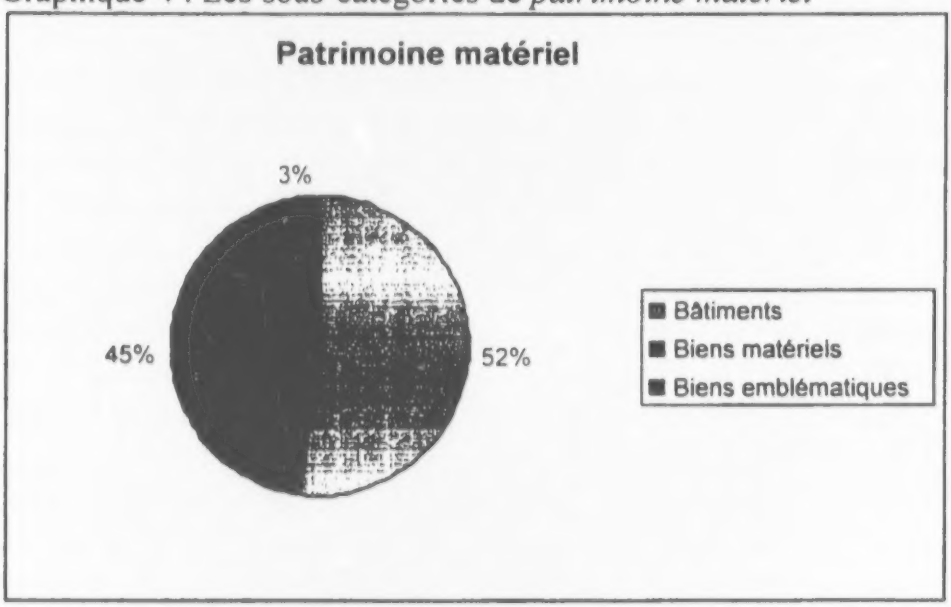


Comme présenté dans le graphique 3, la catégorie *patrimoine matériel* se démarque nettement. Elle comprend 48 % des éléments de réponse. La catégorie *patrimoine immatériel* a aussi obtenu un important pourcentage des réponses, soit 29 %. Les 6 autres catégories se partagent inégalement les derniers 23 %.

Les éléments constitutifs de la représentation du patrimoine suggérée par les répondants sont en majorité de type matériel ou immatériel. Les éléments immatériels se définissent comme des éléments purement symboliques, sans aucune forme de matérialité, tels les valeurs, l'art (littérature, cinéma, peinture, musique, etc.), la langue ou la religion. Pour l'ensemble des gens interrogés, le patrimoine serait mieux représenté par quelque chose de palpable ou de symbolique que par une activité ou un trait génétique, par exemple.

Les catégories *patrimoine matériel* et *patrimoine immatériel* ont été subdivisées, les résultats de cette subdivision sont présentés dans les graphiques 4 et 5.

Graphique 4 : Les sous-catégories de *patrimoine matériel*

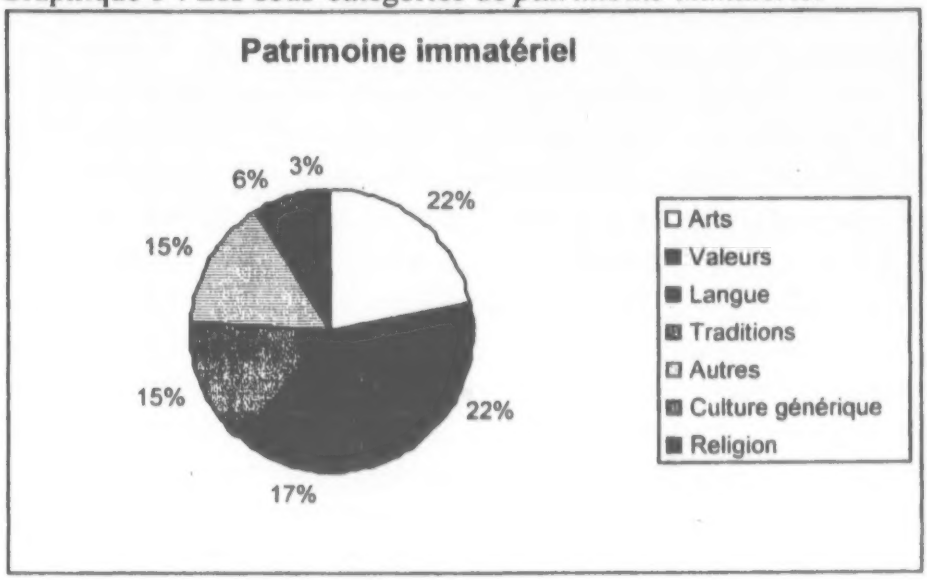


Dans la catégorie *patrimoine matériel*, deux sous-catégories se démarquent : *bâtiments* et *biens matériels*. Les bâtiments constituent 52 % des réponses et les biens matériels 45 %. Les biens matériels emblématiques n'en représentent que 3 %, ce qui signifie que les répondants n'ont pas une vision clichée du patrimoine.

La sous-catégorie *biens matériels* aurait pu être subdivisée davantage. Par exemple, une distinction entre biens mobiliers et biens personnels aurait pu être effectuée. Selon cette catégorisation, des réponses comme « armoire » ou « divan » auraient été classées dans *biens mobiliers*, alors que « bague » ou « photos » l'auraient été sous *biens personnels*. On aurait pu aussi distinguer les biens reçus des biens acquis, les biens personnels des biens collectifs, les biens qui existent dans la permanence du temps des biens éphémères, etc. Mais une telle catégorisation aurait comporté beaucoup d'ambiguïtés et plusieurs catégories se seraient chevauchées. De plus, les

réponses en elles-mêmes étaient souvent trop peu précises pour effectuer une telle catégorisation, ce qui aurait impliqué, dans la majorité des cas, beaucoup d'interprétation de la part des analystes.

Graphique 5 : Les sous-catégories de *patrimoine immatériel*



La catégorie *patrimoine immatériel* se subdivise selon les sept sous-catégories suivantes : *arts*, *valeurs*, *langue*, *traditions*, *autres*, *culture* (sens générique), *religion*. Comme le montre le

graphique 5, les sous-catégories *arts* et *valeurs* prédominent. Chacune représente à elle seule 22 % des réponses.

Au Québec, la langue et la religion sont traditionnellement des éléments constitutifs de l'identité. En conséquence, il est intéressant de constater que si la langue a obtenu 17 % des réponses, la religion n'en a obtenu que 3 %[1]. Il semble que les gens associent davantage le patrimoine à certaines représentations religieuses (églises, art religieux) qu'à la religion elle-même[2].

Réflexions sur la définition du patrimoine

Nous avons cherché à savoir si ce que nous avions observé lors de cette enquête correspondait, du moins en partie, aux définitions de *patrimoine* que nous donnent certains ouvrages de référence. Des liens ont été établis entre les éléments de réponse et les définitions du mot *patrimoine* proposées par le *Nouveau Petit Robert* (1993) et le *Dictionnaire Québécois d'Aujourd'hui* (1993).

D'abord, le *Nouveau Petit Robert* (NPR) définit le mot *patrimoine* de la façon suivante :

> *1 Biens de famille, biens que l'on a hérités de ses ascendants. 2 DROIT Ensemble des biens corporels et incorporels et des créances nettes d'une personne (physique ou morale) ou d'un groupe de personnes, à une date donnée. 3 Ce qui est considéré comme un bien propre comme une propriété transmise par les ancêtres. 4 BIOL. L'ensemble des caractères hérités* (1993).

[1]. L'âge moyen des répondants peut expliquer en partie la faible représentation de la religion.

[2]. Cela vaut aussi pour les membres de communautés religieuses qui ont répondu au questionnaire.

La définition du *Dictionnaire Québécois d'Aujourd'hui* (DQA) présente de nombreuses similitudes avec celle du NPR. Le mot *patrimoine* y est défini comme suit :

> *1| Biens de familles, biens que l'on a hérités de ses père et mère. 2| Ensemble des réalisations considérées comme une propriété transmise par les ancêtres* (1993)

Notons qu'aucun élément de définition du patrimoine proposé par ces ouvrages ne fait référence aux concepts de souvenir et d'identité. Les deux fonctions du patrimoine le plus fréquemment nommées par les répondants ne correspondent pas aux éléments des définitions du patrimoine proposées dans les dictionnaires. Or, pour les répondants, le patrimoine est un agent mémoriel et identitaire important, il sert de référence à un peuple, une famille, etc.

De plus, il est intéressant de constater que le patrimoine devient un outil de réflexion puisqu'il sert à *projeter* et *comprendre*. En effet, 17 % des répondants voient dans le patrimoine un moyen permettant d'analyser un contexte (généralement le passé) pour en comprendre ou prévoir un autre (le présent ou le futur, quelquefois les deux).

Ces utilisations particulières du patrimoine ne se retrouvent jamais, ni de près ni de loin, dans les ouvrages de référence.

Quelles explications pouvons-nous trouver à ce constat ? Assurément, la vision dictionnairique du patrimoine demeure en général plus près du sens « léguer un bien en héritage » ou, du moins, est plus restreinte que les réponses que nous avons obtenues. Toutes les fonctions du patrimoine proposées par les répondants sont absentes des définitions. Malgré l'idée de « legs » que l'on retrouve dans les définitions, la catégorie *transmettre* ne s'y trouve pas plus représentée que les autres catégories. On parle du patrimoine comme quelque chose de transmis, et non pas comme quelque chose qui sert à transmettre. Nous sommes, en conséquence, contraints de voir

que les définitions des dictionnaires ne véhiculent pas les fonctions du patrimoine.

Nous pouvons établir un second contraste entre les résultats de l'enquête et les définitions des dictionnaires. Nos résultats dépassent largement le simple cadre individuel et familial. L'identité collective (63,1 %) représente presque le double des réponses rattachées à l'identité individuelle (36,9 %). De plus, le souvenir des choses (64,3 %) devance amplement le souvenir des gens (26,2 %) ou des origines (13,1 %). Comment expliquer cela ? L'idée du patrimoine, du moins au Québec, est peut-être en évolution. Peut-être reflète-t-elle une tradition différente de celle de la France en ce qui a trait à la transmission des biens et des valeurs.

Les exemples donnés par les répondants iraient aussi en ce sens. Les personnes interrogées ont mentionné beaucoup de biens de famille, par exemple « l'horloge de mon grand-père », « une vieille armoire qui appartient à ma famille depuis très longtemps », mais ils ont aussi nommé un nombre important de biens collectifs, comme « les églises » et « le Château Frontenac ». De même, le NPR et le DQA définissent le patrimoine comme un bien transmissible. Cependant, les répondants considèrent aussi le patrimoine comme un bien non transmissible puisqu'ils ont donné des réponses comme « les églises » et « les Plaines d'Abraham ».

En somme, selon les résultats que nous avons obtenus, le sens premier de *patrimoine* est beaucoup plus étendu que celui répertorié dans le NPR ou le DQA. En conséquence, nous suggérons une définition plus adéquate du patrimoine, tel qu'il est perçu par les gens d'ici, aujourd'hui :

> *Biens, valeurs, coutumes, savoirs transmis à travers le temps, partagés par une collectivité ou une famille, et considérés comme une composante de son identité.*

Bibliographie

BOULANGER, J.-C. (1997), *Lexicographie et dictionnaires*, notes de cours FRN-20517, Québec, Université Laval.

La presse québécoise vue à travers le traitement automatique de ses particularités lexicales

Anne Dister
Université de Liège - Université Paris 7

Résumé

Les logiciels de traitement automatique conçus pour le français posent un certain nombre de problèmes dès lors qu'il s'agit de travailler sur des textes du français hors de France. Le système Intex que nous utilisons, bien que pourvu de dictionnaires à large couverture, ne peut traiter de manière totalement satisfaisante des textes provenant de la Suisse, du Québec ou encore du Sénégal.

Dans la présente étude, nous proposons d'analyser les particularités lexicales d'un corpus de textes québécois qui ne sont pas reconnues par notre dictionnaire électronique des mots simples du français standard.

Le journal *Le Devoir*

Le corpus sur lequel nous travaillons est constitué d'articles parus dans le journal *Le Devoir* sur une période d'environ trois mois, soit de fin 1997 à début 1998. Cela correspond approximativement à 1 045 169 mots simples, ou 1 500 pages.

Pour analyser les données, nous utilisons le logiciel Intex[1], environnement de développement linguistique qui comprend des dictionnaires à large couverture (de mots simples, mots composés, noms propres, abréviations, etc.), lequel permet de travailler sur de très gros corpus.

---

[1] Pour une description détaillée du système, voir Silberztein (1993).

Nous procédons de la manière suivante : nous appliquons sur le texte le dictionnaire des mots simples (qui comprend actuellement près de 900 000 entrées). Tous les mots reconnus par le dictionnaire sélectionné apparaissent dans une case avec leur analyse morphosyntaxique ; les mots non reconnus figurent quant à eux dans une case de mots inconnus. Max Silberztein (1995) a calculé que près de 80 % des mots inconnus étaient des noms propres. Puisque c'est cette case des mots inconnus qui nous intéresse et afin de ne pas encombrer inutilement l'analyse de celle-ci, nous avons choisi d'appliquer au texte un automate qui analyse automatiquement comme des noms propres les formes non reconnues par le dictionnaire des mots simples qui commencent par une majuscule. Sur les 11 031 formes non reconnues restent alors 2 561 formes inconnues. C'est sur ces 2 561 formes que nous allons centrer notre propos, en voyant en quoi elles sont ou non spécifiques d'un corpus du français du Québec.

Les mots mal orthographiés : coquilles et fautes d'orthographe

Un grand nombre de mots de notre corpus ne sont pas reconnus par les dictionnaires que nous avons appliqués[2] car ils sont mal orthographiés. Ce type de problème apparaît peu ou prou dans tous les types de corpus analysés[3]. Dans *Le Devoir*, les principales coquilles se répartissent comme suit :

- une lettre manque : *juilet*, *poucentage*, *quicaillerie*, *suveiller*, *sytème*, *tranfuge*, *vendedi*, *cinmatographique*, etc.
- une lettre est en trop : *juilllet*, *structrure*, *vàioloniste*, etc.

[2] Outre le dictionnaire des mots simples, nous avons également appliqué un dictionnaire de sigles et d'abréviations.

[3] Pour une étude comparable sur une année du journal *Le Monde*, voir Mathieu *et al.* (1998).

- une lettre est mise pour une autre : *babages*, *chanceuce*, *movembre*, *ordinadeurs*, *tinanesque*, etc.
- deux lettres sont interverties : *mathématqiues*, *vérétan*, etc.
- l'accent manque : *millenaire*, *dégats*, etc.
- le mauvais accent est utilisé : *deuxiéme*, etc.
- l'espace manque entre deux mots qui sont alors soudés : *ordonneà*, *seuljuge*, *desVêpres*, etc.

Il y a aussi les erreurs qui s'apparentent plus à des fautes d'orthographe qu'à des coquilles. On peut citer, entre autres : *ortographe*, *floppée*, *dangeureusement*, *dangeureux*, *plaidoieries*, *exhorbitantes*, *succédés*, *participés*, *maronniers*, *ambiguité*, *rafinée*, etc.

L'orthographe réformée

On constate, à observer attentivement les mots non reconnus, qu'un certain nombre d'entre eux qui pourraient a priori être considérés comme mal orthographiés participent en fait des rectifications orthographiques proposées par le Conseil supérieur de la langue française en 1990.

Cette réforme, qui visait des « aménagements » afin de supprimer de l'orthographe française « un certain nombre d'anomalies », ne s'est pas véritablement imposée, même si certaines tentatives continuent de se manifester pour la promouvoir[4].

Néanmoins, et nous pouvons le constater dans notre corpus, de nombreuses occurrences de mots orthographiés selon les recommandations de la réforme sont quotidiennement utilisées dans les textes.

---

[4] Citons par exemple la réédition récente au Seuil du livre dirigé par Jacques Chaurand, *Nouvelle histoire de la langue française*.

Ainsi, nous trouvons dans *Le Devoir* des formes qui résultent des propositions suivantes[5] :

- Suppression de l'accent circonflexe sur le *i* et le *u*.

On trouve les formes : *connaît*, *connaître*, *apparaît*, *paraît*, *comparait*, *accroit*, *indument*, *croutes*, *ile* et *entraineur*.

L'exemple suivant est particulièrement intéressant, puisque l'auteur y considère explicitement cette forme comme une faute d'orthographe, la transcrivant, avec d'autres, entre guillemets :

> *Crawford parle pendant 30 secondes en anglais, et on sous-titre : « neufs enfants », « on grandi », « ligue national », « entraineur ». Il y a du Bescherelle là-dessous...*

- Ajout ou changement d'un accent qui rend la forme graphique conforme à la prononciation. C'est le cas de *artéfacts* ou *règlementation*.

La conjugaison de quelque 240 verbes est affectée par cette rectification[6]. Ainsi, le futur et le conditionnel de verbes comme *régler* n'est plus *réglerai / réglerais* mais *règlerai / règlerais*. Parmi les verbes touchés par cette modification on peut citer *abréger*, *accéder*, *accélérer*, *acérer*, *assécher*, *assiéger*, *avérer*, etc.

Dans *Le Devoir*, nous avons rencontré les formes *règlera*, *règlerait* et *complèteront*.

- Suppression de lettres superflues : *asseoir* devient, comme nous l'avons trouvé, *assoir*.

---

[5] Nous résumons ici. Pour un panorama complet des « règles » proposées, de leur justification ainsi que des formes concernées voir Goosse (1991).

[6] Après flexion, cela donne 2 668 entrées.

- Le pluriel des noms étrangers se fait sur la base des règles morphologiques et phonétiques du français. *Jazzman* au pluriel est ainsi *jazzmans* et non *jazzmen* comme en anglais. *Pizzicato* au pluriel n'est plus invariable mais prend un *s* : *pizzicatos*, comme soprano qui s'écrit au pluriel *sopranos*[7].

Toutes les formes que nous venons de citer dans ce paragraphe étaient absentes du dictionnaire des mots simples. Elles posaient un problème pour l'analyse, mais sans doute moins complexe que le problème posé par l'exemple suivant :

> *Depuis le décès du cinéaste en juin 1982, à l'âge de 37 ans, l'effet Fassbinder s'est partiellement dissout en Amérique.*

On trouve ici la forme *dissout* à la place de *dissous*. Or, *dissout* est répertorié dans le dictionnaire des mots simples d'Intex uniquement comme un indicatif présent de la troisième personne du singulier du verbe *dissoudre*[8]. Il est donc analysé comme tel, ce qui est ici erroné.

Pour résoudre ce problème général des mots non reconnus participant des recommandations de 1990, nous avons créé un dictionnaire électronique de l'orthographe réformée (delafmOR). Pour les mots simples, il comprend dans sa forme actuelle près de 9 000 entrées. Il adopte le format et les codes des autres dictionnaires du système Intex : une ligne correspond à une entrée lexicale qui se compose comme suit :

- la forme telle qu'on peut la trouver dans le texte ;
- une virgule ;

---

[7] Notons à cet égard, pour relever (s'il était encore nécessaire de le faire) l'incohérence de l'orthographe française, que *mezzo-soprano* au pluriel, dans l'orthographe courante, prenait quant à lui déjà l'*s* au pluriel (*des mezzo-sopranos*).

[8] La réforme propose d'écrire les participes passés *dissous* et *absous*, *dissout* et *absout*.

- le lemme (c'est-à-dire la forme canonique, telle qu'on la trouve en entrée dans un dictionnaire courant) ;
- un point ;
- la catégorie grammaticale ;
- un double point ;
- les informations flexionnelles (genre et nombre pour les noms et les adjectifs, conjugaison pour les verbes).

Voici un extrait des premières entrées du dictionnaire delafmOR[9] :

abime,abime.N:ms
abimes,abime.N:mp
abrègement,abrègement.N:ms
abrègements,abrègement.N:mp
absout,absoudre.V:Kms
absouts,absoudre.V:Kmp
addenda,addenda.N:ms
addendas,addenda.N:mp
affèterie,affèterie.N:fs
affèteries,affèterie.N:fp
affut,affut.N:ms
affutage,affutage.N:ms
affutages,affutage.N:mp
affuteur,affuteur.N:ms
affuteurs,affuteur.N:mp
affuteuse,affuteur.N:fs
affuteuse,affuteuse.N:fs
affuteuses,affuteur.N:fp
affuteuses,affuteuse.N:fp
affutiaux,affutiaux.N:mp
affuts,affut.N:mp
aigüe,aigu.A+d:fs
aigües,aigu.A+d:fs
ainé,ainé.A+d:ms
ainé,ainé.N:ms
ainée,ainé.A+d:fs
ainée,ainé.N:fs
ainées,ainé.A+d:fp
ainées,ainé.N:fp
ainées,ainés.N:fp
ainés,ainé.A+d:mp
ainés,ainé.N:mp

L'utilisation de notre dictionnaire de l'orthographe réformée, parallèlement au dictionnaire des mots simples, permet de reconnaître tous les mots cités dans cette section et, pour la forme *dissout*, de donner la double analyse suivante :

dissout,dissoudre.V:P3s

[9] Nous présentons les entrées ici sur deux colonnes.

dissout,dissoudre.V:Kms[10]

Le problème est évidemment que cette double possibilité d'analyse sera fournie également pour les occurrences où *dissout* est bien un indicatif présent.

La féminisation des noms de métier, fonction, grade ou titre

En 1979 paraissait dans la *Gazette officielle du Québec* un avis recommandant de féminiser les titres. Depuis de nombreuses années, des formes comme *l'auteure*, *la professeure* ou *la juge* sont courantes en français du Québec.

La Suisse, en 1988, et la Belgique, en 1993, ont également pris des mesures institutionnelles visant à encourager l'utilisation des formes féminines.

Dans *Le Devoir*, il semble que toutes les professions exercées par des femmes soient systématiquement notées sous leur forme féminine. On trouve à de nombreuses reprises les formes, inexistantes en français de France ou de Belgique, *auteure*, *professeure*, *compositeure*, *metteure en scène*, *Procureuse générale*, *réviseuse*, etc.

Par ailleurs, certains noms de métier sont épicènes et seuls les déterminants et les adjectifs prennent la marque du féminin : on trouve ainsi très fréquemment dans *Le Devoir*, *la juge* et *la ministre*.

Ainsi, outre les nouvelles formes qu'il faut répertorier et encoder, il s'avère nécessaire de compléter dans le dictionnaire les

---

[10] V :P3s = verbe à la troisième personne du singulier de l'indicatif présent ;
V:Kms = Verbe, au participe passé masculin singulier.

informations morphologiques associées aux mots épicènes et, par exemple, transformer l'entrée « juge,.N:ms » en « juge,.N:ms:fs ».

Ce travail a été fait systématiquement et constitue le *Dictionnaire des noms de profession* qui comprend 4 190 formes fléchies[11]. Néanmoins, se basant sur le guide publié par la communauté française de Belgique, il n'inclut pas, à l'heure actuelle, les formes féminisées typiquement québécoises mentionnées plus haut.

Les québécismes

Comme on pouvait s'y attendre, un grand nombre de formes non reconnues dans *Le Devoir* sont en fait des formes propres au français du Québec. On peut citer entre autres : *cenne*, *s'enfarge*, *s'enfargeant*, *enfirouaper*, *frimée*, *tounes*, *fun*, *guidounes*, *chum*, *tabletter*, *se garrochant*, *baloune*, *jasette*, *diseux*, *rapaillé*, *ratoureux*, *ratoureuse*, *recherchiste*, *robineux*, *bebitte*, *bibitte*, *chialeux*, *patenteux*, *parlable*, *pelleteux*, *perchaude*, *quêteux*, *saumonnier*, *tripeux*, *tripeuse*, *téteux*, *québécisants*, *joualisants*, etc.

Il faut noter que de nombreuses occurrences de québécismes populaires interviennent dans du discours rapporté. On a notamment :

> *Elle cherche un frigo, pose des questions à tout le monde, déclare que « tout a l'air ben plus l'fun de ce bord-citte »* [...].
>
> *« On ne joue pas les vierges offensées », dit Benoît, « sans être guidounes non plus », de préciser Jean-Pierre.*

[11] Il est téléchargeable à l'adresse suivante : http://www.ladl.jussieu.fr/tools/tools.html#profession

> *Fin psychologue, « il faut leur faire accepter la douleur et leur faire la jasette », dit-il.*

À cette liste, s'ajoutent tous les adjectifs dérivés de noms de lieux : *abitibien, abénaquaise, albertaines, kongais, kongaise, kongaises, laurentien, laurentienne, lavalloise, manitobaines, rimouskois, saskatchewanais*, etc.

La constitution d'un *Dictionnaire électronique des formes simples en français du Québec* (DELQUES) est en cours à l'UQAM. Il contient actuellement 87 505 entrées lexicales et, par le fait même, permet d'étendre la couverture lexicale des corpus québécois. En voici un extrait[12] :

```
cenne,cenne.N21+z1+Q:fs/(Q)
cennes,cenne.N21+z1+Q:fp/(Q)
[...]
gratteau,gratteau.N3+z3+Q:ms/(Q)
gratteaux,gratteau.N3+z3+Q:mp/(Q)
gratteuse,gratteux.A63+z1+Q:fs/(Q)
gratteuse,gratteux.N63+z1+Q:fs/(Q)
gratteuses,gratteux.A63+z1+Q:fp/(Q)
gratteuses,gratteux.N63+z1+Q:fp/(Q)
gratteux,gratteux.A63+z1+Q:mp/(Q)
gratteux,gratteux.A63+z1+Q:ms/(Q)
gratteux,gratteux.N2+z1+Q:ms:mp/(Q)
gratteux,gratteux.N63+z1+Q:mp/(Q)
gratteux,gratteux.N63+z1+Q:ms/(Q)
```

Néanmoins, la constitution d'un tel dictionnaire ne règle pas le problème trivial des faux-amis :

> [...] *ce n'est pas le dépanneur du coin mais Coca-Cola, McDonald's, IBM, Kodak qui allongent au bas mot une trentaine de millions* [...].

---

[12] Le nombre qui suit la catégorie grammaticale (ici *A* ou *N*) est le code de flexion, *z* le numéro de couche lexicale. (*Q*) indique qu'il s'agit du français du Québec.

Dans cet exemple, le sens de *dépanneur* est totalement inconnu en France ou en Belgique.

Une autre version du dictionnaire est également en chantier, qui reprend notamment des marques sémantiques et donne, lorsqu'il y a lieu, un équivalent sémantique (entre angles) qui permet de distinguer les faux-amis. En voici un extrait[13], avec notamment les entrées pour *dépanneur* :

argenté,A32+z3+Q/LAP/BER/GLO/<riche>
argenteries,N22P+Conc+z2+Q/GLO/<argenterie>
argents,N2P+Abst+z1+Q/GLO/DUL/BER/<fonds;valeurs>
argnée,N31+Conc+[Obj]+z3+Q/BER/<crochet servant à remonter les seaux du puits>
argot,N1+Conc+Pc+z1+Q/BER/<sabot>
argotage,N1+Abst+z3+Q/BER/=ergotage
arguer,V3+t+32R3+~cause+z3+Q/GLO/<plaider>
arguette,N21+Conc+[Obj]+z3+Q/BER<oreiller>
argument,N1+Abst+z2+Q/BER/GLO/<dispute>
aria,N1+Abst+z1+Q/DUL/SEU/LAP/DQA/<embarras>
aria,N1+Abst+z1+Q/DUL/BER/SEU/DQA/GLO/<gâchis;fouillis>
aria,N1+Conc+z1+Q/DUL/BER/DQA/GLO/<attirail>
aria,N1+Hum+z1+Q/DQA/<personne qui est un fardeau>
[...]
dépanneur,N1+Conc+[Lieu]+z1+Q/DUL/SEU/<épicerie du coin>/=accomodation
dépanneur,N35+Hum+z2+Q/DUL/SEU/<tenancier de dépanneur>
dépanneur,N1+Conc+[Obj]+z2+Q/DUL/<grille pour déplacer un véhicule dans la neige>

[13] Conc = Concret ; Abst = Abstrait ; DUL = *Dulong* ; GLO = *Glossaire du parler français au Canada.*

Les néologismes

Nous avons rencontré un certain nombre de néologismes qui ne nous semblent pas spécifiques à la « nature québécoise » du corpus.

Dans les deux premiers exemples ci-dessous, l'auteur marque bien l'emploi particulier d'un nouveau terme par la présence de guillemets :

> *Naissance d'une « banlieuelogie ».*
>
> *Qu'on l'aime ou qu'on le déteste, Charles de Gaulle est un personnage « hénaurme ».*
>
> *Ils sont fous ces muséologues! On peut tout de même séparer le bon grain muséologique de l'ivraie muséomaniaque dans ce lot, en se concentrant d'abord sur les établissements consacrés à l'histoire ou au patrimoine* [...].
>
> *Passons sur les manifestations coutumières de « l'hommerie » que sont le vandalisme, le vol et le non-respect totalement incivique des consignes de parcimonie dans l'utilisation du courant, ce n'est pas de la semaine dernière qu'on descend du singe.*

On notera dans l'exemple suivant l'utilisation de la forme féminisée *acteure*, création de la féministe Pol Pelletier :

> *Il est alors mûr pour autre chose que va lui offrir Pol Pelletier, elle-même sans pareil, fondatrice de l'avant-gardiste Théâtre expérimental, papesse du très féministe Théâtre expérimental des femmes (TEF), qui dirige maintenant le Dojo pour « acteures ».*

Pour terminer, nous pouvons également citer les néologismes créés grâce à des préfixes productifs, et qui ne sont pas spécifiques au français du Québec :

anticascades, *anticoncentration*, *antifraude* ;

coréalisation, *coscénariser*, *coscénarisé*, *cyberattaques*, *cyberclasse*, *cybercommerce*, *cybercours*, *cyberculture*, *cyberespace*, *cybermaniaques*, *cybermartyr*, *cybernautes*, *cyberspace*, *cyberspatial*, *cyberspatiale* ;

hyperlien, *hyperliens*, *hyperlucratif*, *hypermédias*, *hypersophistiquée*, *hyperventilé* ;

mégacommandes, *mégacommunauté*, *mégacomplexes*, *mégacontrat*, *mégafichier*, *mégafichiers*, *mégafusion*, *mégaoctects*, *mégaphénomène*, *mégashows*, *mégasite*, *mégastars*, *mégasuccès*, *mégasurface*, *mégatransaction*, *megawatts*.

Intex n'a actuellement pas de procédure d'analyse automatique de ces néologismes. Ce problème des néologismes est donc pour nous le même que pour les lexicographes : faut-il entrer tous ces mots, peut-être création d'un jour, dans le dictionnaire ?

Conclusion

Quel que soit le corpus qu'il analyse, le chercheur en traitement automatique des langues est confronté à des problèmes invariants : les coquilles et fautes d'orthographe, les termes spécifiques au corpus utilisé et qui ne sont pas présents dans les dictionnaires du système. Nous avons vu les solutions que le système Intex peut apporter. Deux nouveaux dictionnaires, valables pour toutes les variétés de français, permet de reconnaître un certain nombre d'occurrences jusque-là ignorées : le dictionnaire des noms de profession et le dictionnaire de l'orthographe réformée. Par ailleurs,

 BEST COPY AVAILABLE

pour les corpus québécois, l'application du *Dictionnaire électronique des formes simples en français du Québec* (DELQUES) améliore considérablement la couverture lexicale.

Bibliographie

BOULANGER, J.-C. (dir.) (1992), *Dictionnaire québécois d'aujourd'hui*, Saint-Laurent (Québec), DICOROBERT INC.

CHAURAND, J. (dir.) (1999), *Nouvelle histoire de la langue française*, Paris, Seuil.

DISTER, A. (1999), « Dictionnaires électroniques de l'orthographe réformée », *Actes du XVIIIe colloque Lexiques-grammaires comparés, Salerne (6-9 octobre 1999)*, Salerne. À paraître.

GOOSSE, A. (1991), *La « nouvelle » orthographe*, Paris / Louvain-la-Neuve, Duculot.

Guide de féminisation des noms de profession, Communauté française de Belgique.

LABELLE, J. (1994), *Dictionnaire électronique des formes simples en français du Québec*, DELQUES V1.0, Rapport de recherche n° 9, Montréal, GRFL, UQAM.

LABELLE, J. (1996), « Linguistique comparée et dictionnaires électroniques et INTEX », *Actes des premières journées INTEX*, Paris, Université Paris 7.

MATHIEU, Y. Y., G. Gross, C. Fouqueré (dir.) (1998), « Vers une extraction automatique des néologismes », *Cahiers de Lexicologie*, n° 72, p. 199-208.

POIRIER, C. (dir.) (1998), *Dictionnaire historique du français québécois*, Sainte-Foy, Les Presses de l'Université Laval.

SILBERZTEIN, M. (1993), *Dictionnaires électroniques et analyse automatique de textes, Le système INTEX*, Paris, Masson.

SILBERZTEIN, M. (1995), « Dictionnaires électroniques et comptage de mots », *JADT 1995*, Rome, vol. 1, p. 93-101.

Les expressions partitives en français et en espagnol

Claudine Leullier
Université Laval

Le français et l'espagnol, bien qu'elles soient deux langues romanes, comportent des différences importantes en ce qui a trait aux expressions partitives. En français, contrairement à ce qui se produit en espagnol, la structure partitive est généralement composée d'une préposition suivie d'un article (je veux du vin / quiero vino). D'après notre hypothèse, cette différence s'explique, notamment, par le fait que l'expression partitive est étroitement liée à la présence de quantificateurs et de leur variabilité ou non. Le manque de concordance entre le quantificateur et le SN – ce qui se produit fréquemment en français – entraînerait la présence de la préposition *de* et de l'article. En espagnol, au contraire, il y a concordance entre le quantificateur et le SN dans la plupart des cas, d'où l'absence de la préposition *de* et de l'article. D'après ce raisonnement, on peut conclure que les expressions françaises du type *je veux du vin* contiendraient un quantificateur vide sans concordance et que les expressions espagnoles, telles que *quiero vino*, contiendraient un quantificateur vide avec concordance.

Définition des touts partitifs

Nous définissons l'expression partitive d'une façon sémantique, comme étant une structure qui exprime une partie d'un tout. (Milner, 1978 : 62). Cette structure est composée de deux éléments :

a) Le tout : il peut être un syntagme nominal, voir l'exemple (1), ou un syntagme déterminant, voir l'exemple (2). Ce syntagme peut

dénoter un groupe dénombrable, voir en (2), ou bien une quantité indéterminée d'une substance, voir en (1) :

(1) a. Un peu d'**eau**.
b. Algo de **agua**.
(2) a. Beaucoup des (de + **les**) **femmes** présentes.
b. Muchas de **las mujeres** presentes.

b) La partie du tout : c'est un syntagme déterminant composé soit d'un quantificateur, voir l'exemple (3), et probablement d'un nom vide (Milner, 1978), soit d'un substantif quantitatif, voir l'exemple (4). Ceux-ci quantifient le syntagme nominal ou le syntagme déterminant qui représente le tout et qui est l'ensemble restrictif du quantificateur :

(3) a. J'ai **trop** de devoirs.
b. Tengo **demasiadas** tareas.
(4) a. **Un tas** de terre.
b. **Un puñado** de tierra.

Dans notre définition d'expression partitive, nous incluons des exemples sans quantificateur exprimé :

(5) a. J'ai mangé du pain.
b. He comido pan.
(6) a. Acheter de la farine.
b. Comprar harina.

Nous pouvons les joindre à notre définition car ils sont équivalents sémantiquement à (7) et (8) :

(7) a. J'ai mangé du pain. ⇔ J'ai mangé un peu de pain.
b. He comido pan. ⇔ He comido algo de pan.
(8) a. Acheter de la farine. ⇔ Acheter un peu de farine.
b. Comprar harina. ⇔ Comprar algo de harina.

L'équivalence sémantique des exemples précédents peut s'expliquer en considérant qu'il existe un quantificateur vide, c'est ce que proposent quelques auteurs dont Richard Kayne (1983) et Jacob Hoeksema (1996). Voir l'exemple (9) où Q a une signification équivalente à celle de « un peu ». Q est donc indéfini :

(9) a. J'ai mangé Q du pain.
b. He comido Q pan.

Distribution

L'expression partitive peut se retrouver dans différentes positions tant dans une phrase en français qu'en espagnol. La plupart du temps, on la retrouve dans la position d'un objet direct :

(10) a. Je lis beaucoup de livres.
b. Leo muchos libros.

Parfois on la retrouve dans la position de sujet :

(11) a. Du pain nous suffit. (Grévisse, 1969)
b. Nos bastará pan.

On retrouve quelquefois des expressions partitives comme complément de certaines prépositions :

(12) a. J'ai fait le plancher avec de la céramique.
b. Hice el piso con cerámica.

Le fait que les expressions partitives du français et les syntagmes sans déterminant de l'espagnol aient pareille distribution indique qu'il y a une relation syntagmatique entre ces deux constructions.

Problème à l'étude

Le problème à l'étude est d'expliquer les différences entre le français et l'espagnol, c'est-à-dire expliquer pourquoi en français on retrouve souvent la présence de la préposition *de* et d'un article quand il n'y a pas de quantificateur alors qu'en espagnol on ne retrouve la plupart du temps aucun de ces deux indices :

(13) a. Je veux [$_{SD}$ Q [$_{SP}$ de [$_{SD}$ le vin.]]]
b. Quiero [$_{SD}$ Q [$_{SN}$ vino.]]
(14) a. Il boit [$_{SD}$ Q [$_{SP}$ de [$_{SD}$ l'eau.]]]
b. Bebe [$_{SD}$ Q [$_{SN}$ agua.]]

Hypothèse

Tel qu'expliqué dans notre définition de l'expression partitive, chacune de ces structures a un quantificateur explicite ou implicite. Notre hypothèse pour tenter d'expliquer les différences entre ces deux langues est que la variabilité ou l'invariabilité de ce quantificateur amène une construction particulière de partitive. De cette façon, il y aurait une relation entre l'existence de concordance ou non et l'apparition de la préposition *de* ou non :

(15) a. J'ai trop de devoirs.
b. Tengo demasiadas tareas.

Dans les exemples précédents, *trop* doit se construire avec *de* parce qu'il n'y a pas de concordance et *demasiadas* doit se construire sans *de* parce qu'il y a concordance.

Pour démontrer notre hypothèse, nous commencerons par observer les expressions avec quantificateurs explicites, pour ensuite comparer ces exemples avec ceux qui ont un quantificateur implicite.

Différentes constructions avec quantificateurs explicites

Nous pouvons distinguer six groupes de base pour les constructions partitives avec quantificateurs explicites :

(16) Quantificateur variable, sans *de*, sans article
(17) Quantificateur variable, *de*, article
(18) Quantificateur invariable, *de*, sans article
(19) Quantificateur invariable, *de*, article
(20) Substantif de valeur quantitative, *de*, sans article
(21) Substantif de valeur quantitative, *de*, article

Les exemples (22b) et (25b) appartiennent au groupe (16), (23a et b) au groupe (17), (22a) et (25a) au groupe (18), (24a et b) au groupe (19), (26a et b) au groupe (20) et les numéros (27a et b) au groupe (21).

(22) a. J'ai trop de problèmes.
b. Tengo demasiados problemas.
(23) a. J'ai fait quelques-uns des voyages offerts par la compagnie.
b. Hice algunos de los viajes ofrecidos por la compañía.
(24) a. J'ai utilisé peu de la terre que tu m'as donnée.
b. Usé algo de la tierra que me diste.
(25) a. Il vend beaucoup de fruits.
b. Vende muchas frutas.
(26) a. J'ai acheté un litre de lait.
b. Compré un litro de leche.
(27) a. J'ai cuisiné un kilo des patates que tu as cultivées.
b. Cociné un kilo de las papas que cultivaste.

Il est à noter qu'aucun tout partitif avec quantificateur variable et sans article a dans sa structure la préposition *de* et que tous les groupes avec quantificateur invariable sont structurés avec la

préposition *de*. Ces deux constatations appuient directement notre hypothèse.

Quantificateurs variables

Premièrement, les constructions avec quantificateurs variables, sans *de* et sans article :

(28) a. Quelques enfants patinent.
b. Algunos niños patinan.
(29) a. Je lis plusieurs livres chaque mois.
b. Leo muchos libros cada mes.

Ces constructions avec quantificateurs variables ne semblent pas partitives, si on se souvient de notre définition, puisqu'à première vue il n'est pas évident de déterminer ce que représente le tout, ou même s'il existe un tout. Cependant, nous considérons que ces expressions sont partitives et que le tout est implicite.

Deuxièmement, les constructions avec quantificateurs variables, avec *de* et avec article. Faisons un parallèle avec la structure précédente :

(30) a. Aucune femme ne croit que c'est important.
b. Aucune des femmes du groupe ne croit que c'est important.

(31) a. Ninguna mujer cree que es importante.
b. Ninguna de las mujeres del grupo cree que es importante.
(32) a. J'ai lu plusieurs livres.
b. J'ai lu plusieurs des livres de l'auteur que tu m'as parlé.
(33) a. He leído muchos libros.

b. He leído muchos de los libros del autor que tú me hablaste.

Nous pouvons observer qu'au numéro (30a) le tout n'est pas exprimé, il n'y a donc ni article ni préposition *de*. Au numéro (30b), le tout est exprimé et est défini, il y a donc un article et la préposition *de*. La même chose se produit entre les numéros (31a et b), (32a et b) et (33a et b).

Quantificateurs invariables

À présent, analysons les constructions avec quantificateurs invariables, avec *de* et sans article et les constructions invariables avec *de* et un article. Comparons ces deux structures :

(34) a. Pedro comió algo de pan.
b. Pedro comió algo del pan que hiciste.
(35) a. Tu as assez de problèmes.
b. Tu as assez des problèmes de ton fils.
(36) a. Ana no bebió nada de limonada en todo el día.
b. Ana no bebió nada de la limonada que preparaste.
(37) a. J'ai mangé beaucoup de bonbons.
b. J'ai mangé beaucoup des bonbons que ta mère m'a donnés.

Ces parallèles nous montrent qu'un même quantificateur invariable peut avoir deux constructions différentes. Avec ces exemples nous observons que si le tout est défini, nous avons l'article, et si le tout est indéfini, il n'y a pas d'article.

Une différence entre les quantificateurs invariables du français et de l'espagnol est que des mots comme *beaucoup* ou *assez* se différencient des mots espagnols comme *algo* ou *nada*, parce que ces derniers ne peuvent se construire avec des SNs pluriels :

(38) a. J'ai assez de farine.
b. J'ai assez de livres.
(39) a. Tengo algo de pan.
b. * Tengo algo de frutas.

Substantifs quantifiables

Maintenant, les constructions avec substantifs quantitatifs. Elles ont la même structure que celle incluant un quantificateur invariable, c'est-à-dire avec *de* et avec l'article quand le tout est défini :

(40) a. Un tas de terre.
b. Un montón de tierra.
(41) a. Un kilo des patates de ma tante.
b. Un kilo de las papas de mi tía.
(42) a. Un litre de lait.
b. Un litro de leche.
(43) a. Une douzaine d'œufs.
b. Una docena de huevos.

Par contre, ces expressions ne sont pas des quantificateurs mais des noms, car autant *tas* que *litre* que *kilo* et que *douzaine* sont clairement des noms. Nous pouvons prouver cette catégorie par le fait qu'ils peuvent être modifiés, contrairement à tous les quantificateurs précédents :

(44) a. Un grand tas de terre.
b. Un gran montón de tierra.
(45) a. Un demi-kilo de patates.
b. Medio kilo de papas.

Une autre différence avec les quantificateurs est qu'en espagnol les substantifs quantifiables acceptent les syntagmes pluriels, voir les exemples (39a et b) :

(46) Una docena de hombres.

Dans cette catégorie de substantifs il semble adéquat d'inclure *un peu* :

(47) a. Un peu de pain.
b. Un poco de pan.

Cependant, il convient d'observer deux grandes différences :

Premièrement, ce mot ne peut être modifié, il ne s'agit donc pas d'un substantif :

(48) a. * Un bon peu de pain.
b. * Un buen poco de pan.

Deuxièmement, *un peu* ne peut fonctionner avec des SNs pluriels. Comparons le numéro (46) avec le prochain exemple :

(49) * Un poco de hombres.

En conséquence, *un peu* serait un quantificateur même si son article *un* nous laisse croire, à prime abord, que c'est un substantif quantitatif.

Nous avons trouvé d'autres constructions similaires aux substantifs quantitatifs :

(50) a. Una bande de jeunes.
b. Una banda de jóvenes.
(51) a. Une équipe de hockey.
b. Un equipo de hockey.

Nous croyons que ces expressions ne sont pas partitives. Observons que si nous changeons l'objet direct par *en*, en français, nous n'obtenons pas le même résultat qu'avec tous les autres quantificateurs antérieurs et les substantifs quantifiables. Le *en* inclut le substantif au numéro (54) et l'exclut au numéro (56) :

(52) Il a acheté une équipe de hockey.
(53) * Il en a acheté une équipe.
(54) Il en a acheté une.
(55) Il a acheté un kilo de patates.
(56) Il en a acheté un kilo.
(57) * Il en a acheté un.

Revenons maintenant aux constructions avec quantificateurs implicites.

Les expressions sans quantificateur explicite

Nous pouvons maintenant associer les expressions partitives qui ont un quantificateur explicite avec celles qui en ont un implicite. On ne peut pas penser que *des*, *de le*, *de la*, *de les* sont des quantificateurs en eux-mêmes parce que si en français on remplace l'objet direct par *en*, ces mots ne restent pas hors du complément comme dans le cas des quantificateurs.

(58) J'ai acheté des pommes.
(59) * J'en ai acheté des.
(60) J'en ai acheté.

Nous considérons que quand il n'y a pas la préposition *de* comme dans (5b) et (6b), le quantificateur est comme un quantificateur variable. Il semble logique de croire cela car en espagnol il y a plus de quantificateurs variables (sans *de*). Ce serait donc pour cette raison qu'on aurait beaucoup d'exemples sans quantificateurs, sans

de et sans article. Le contraire se produirait en français, c'est-à-dire quand il y a la préposition *de* comme dans (5a) et (6a), cela serait comparable aux expressions avec quantificateur invariable.

Cependant, un phénomène différent se produit en français avec les phrases comme (5a) et (6a) : il y a un article. Dans les quatre constructions que nous avons étudiées avec quantificateurs, celles qui avaient un article étaient celles qui avaient un tout défini, par contre, ici, l'article ne signifie pas que l'expression a un tout défini, au contraire, le tout est indéfini. Pour déterminer ce genre de structure sans quantificateur nous enlevons la préposition *de* comme nous pouvons le voir dans les exemples (64) et (65) :

(64) J'ai acheté de la farine.
(65) J'ai acheté la farine que tu m'as conseillée.

Selon Kayne (1983), ce serait un problème syntaxique. Mais nous n'élaborerons pas davantage sur ce point, puisque ceci dépasse les limites de la présente étude, nous le garderons plutôt pour d'autres recherches éventuelles.

Conclusion

Ce que nous avons voulu montrer dans cet article est qu'en espagnol les expressions partitives sont différentes de celles du français, mais que toutes deux sont liées aux quantificateurs qui peuvent être explicites ou implicites. La variabilité ou l'invariabilité des quantificateurs donne des constructions différentes comme on a pu le constater par l'absence ou la présence de la préposition *de* et d'un article qui, en général, définit le tout. Cette étude est importante non seulement pour la compréhension des constructions partitives du point de vue sémantique et syntaxique, mais aussi pour l'apport qu'elle promet aux études portant sur l'apprentissage d'une langue seconde. Nous en sommes seulement au commencement de cette recherche.

Bibliographie

BARWISE, J., et R. COOPER (1981), « Generalized quantifiers and natural language », *Linguistics and Philosophy* 4, Boston et Dordrecht, D. Reidel Pub., p. 159-219.

BEDEL, J.-M. (1997), *Grammaire de l'espagnol moderne*, Paris, Presses Universitaires de France.

GREVISSE, M. (1965), *Précis de grammaire française*, Paris-Gembloux, Éditions Duculot, p. 74-81.

HOEKSEMA, J. (1996), *Partitives: studies on the syntax and semantics of partitive and related constructions*, Berlin et New York, Mouton de Gruyter.

JONES, M. A. (1996), *Foundations of French Syntax*, Cambridge, Cambridge University, p. 215-244.

KAYNE, R. (1983), *Connectedness and Binary Branching*, Dordrecht, Foris Publications.

LEONETTI J. M. (1992), *El artículo y la referencia*, Madrid, Taurus universitaria.

Éducation bilingue : conséquences sur le comportement linguistique à l'adolescence

Céline Rostiaux
Université Libre de Berlin

La recherche sur le bilinguisme individuel ou familial s'est consacrée particulièrement à l'étude des diverses méthodes d'éducation bilingue et des facteurs d'influence de leur succès respectif ainsi qu'à celle de l'acquisition du langage de l'enfant bilingue (voir, entre autres, Kielhöfer et Jonekeit (1983), Romaine (1989) et Hoffmann (1991)). Ce faisant, les conséquences à moyen et à long terme du choix de ces méthodes n'ont pas ou peu été considérées. Or, il est important de ne plus négliger cet aspect de l'éducation bilingue, car si l'acquisition de la langue par l'individu bilingue est une composante vitale de son rapport à la langue, sa façon de gérer sa bilingualité au quotidien en est une d'importance également. L'étude présentée ici a pour objet le comportement linguistique d'adolescents ayant reçu une éducation bilingue. L'approche sociolinguistique choisie tient compte également d'autres paramètres, notamment psychologiques. Les données utilisées sont issues d'un corpus constitué spécialement pour l'étude, réalisée dans le cadre d'un mémoire de maîtrise à l'Université Libre de Berlin. Cette recherche regroupait une dizaine d'adolescents ayant reçu une éducation bilingue, c'est-à-dire ayant été en contact intensif avec les deux langues avant leur troisième année. Les informateurs avaient entre 12 et 16 ans au moment de l'enquête (janvier-mars 1998) et parlaient l'allemand et le français, langues avec lesquelles ils étaient en contact au quotidien, que ce soit en les parlant eux-mêmes ou, au moins, en les entendant. Ils vivaient tous à Berlin, étudiaient pour la plupart au Lycée Français et étaient issus de familles franco-allemandes ou françaises.

L'objectif de l'étude était d'étudier le rapport entre la nature de l'éducation bilingue reçue par les adolescents et leur comportement, tant en ce qui concerne la production que l'attitude, la perception et l'identité. Les résultats de cette recherche ne se basent naturellement que sur un nombre réduit d'informateurs, de sorte qu'ils ne sont pas généralisables dans l'absolu et se limitent aux informateurs concernés.

La terminologie généralement employée dans le domaine des phénomènes de contact des langues n'étant pas uniforme, il est nécessaire de faire d'abord quelques remarques terminologiques : alternance de langue désigne ici le passage d'une langue à l'autre tant entre deux tours de parole qu'entre deux phrases ou deux propositions. Le terme *emprunt* se réfère à l'insertion d'un seul syntagme de langue A dans un discours en langue B ; il peut s'agir d'un emprunt répandu ou d'un emprunt spontané : ces deux types d'emprunts ont été traités séparément, compte tenu des rôles différents qu'ils semblaient jouer. Le *tag-switching* est l'emploi répété d'une expression de A en B peu restrictive sur le plan syntaxique, telle que *tu sais*, *tu vois*, etc. (Romaine, 1989 : 262). Le terme *interférence* désigne les influences d'une langue sur l'autre sur les plans phonologique, syntaxique ou lexical, dont le locuteur n'est en général pas conscient.

Il est également nécessaire de considérer la recherche sur l'éducation bilingue. Les méthodes d'éducation bilingue les plus répandues et les plus étudiées par les chercheurs sont les suivantes : une personne / une langue, langue de la famille vs. langue de l'environnement dans une famille unilingue, langue de la famille vs. langue de l'environnement dans une famille bilingue, le principe du lieu (chaque langue est parlée dans le pays lui correspondant) ou l'absence de méthode appliquée consciemment (Kielhoefer et Jonekeit, 1982). Mahlstedt (1996) a décrit les facteurs décisifs du succès de l'éducation bilingue, facteurs qui ont été retenus lors du choix des informateurs.

Dans le domaine de la recherche sur les phénomènes de contact des langues, plusieurs approches ont été suivies, et c'est celle de René Appel et Pieter Muysken (1987) qui a été retenue pour la présente étude. Appel et Muysken ont appliqué la catégorisation de Jakobson des fonctions de la langue à l'étude du comportement linguistique bilingue (fonctions référentielle, directive et intégrative, expressive, phatique, métalinguistique et poétique). La fonction référentielle est appliquée par le locuteur lorsqu'il a recours à l'alternance pour trouver le mot juste, celui qui correspond exactement à ce qu'il veut dire. La fonction directive sert en particulier à changer d'interlocuteur en changeant de langue. La fonction expressive permet au locuteur d'exprimer son identité bilingue et biculturelle. La fonction poétique, elle, permet d'utiliser les deux langues afin de jouer avec la langue, de jouer avec les mots. Quant à la fonction métalinguistique, elle permet de commenter une langue dans l'autre. D'autres fonctions sont celles étudiées par Poplack (1989), par exemple, qui se penche sur les citations, commentaires et la recherche du mot juste dans le cadre de son étude des phénomènes de contact des langues.

Remarques méthodologiques

Les adolescents étudiés viennent tous de familles comparables sur le plan sociolinguistique, le critère de variation retenu étant la méthode d'éducation bilingue (incluant principes et stratégies) appliquée par les parents. Chacune des cinq familles étudiées avait appliqué l'une des méthodes citées ci-dessus. La méthodologie consistait en l'observation participante lors de trois séquences organisées avec chacun des adolescents et enregistrées (repas familial, séquence avec les frères et sœurs, sortie avec un / des ami(s) dans un lieu public). Ensuite, chaque adolescent a été interrogé sur sa perception de son propre comportement linguistique et son attitude face aux deux langues, sur le niveau de langue de son entourage, etc. Les parents ont également été

interrogés sur leur appréhension de l'éducation bilingue, du comportement linguistique de leurs enfants, etc.

La langue commune : le « frallemand »

Afin de souligner leur appartenance au groupe linguistique auquel ils aspirent, les adolescents se servent de leur bilingualité comme d'un signe identitaire. Tous disent parler le « frallemand », sans se rendre compte qu'ils ne parlent pas tous le même frallemand, car ils utilisent les éléments du discours bilingue de façon absolument différente, même si tous font à peu près usage de tous ces éléments, comme l'a montré l'analyse des données. Certains se cantonnent à l'usage du français tout en faisant grand emploi d'emprunts spontanés et de *tag-switchings*, alors que d'autres ont à peine recours à l'emprunt mais font un usage très intense de l'alternance de langues.

La langue des frères et sœurs

La langue des frères et sœurs est dans tous les cas la langue forte des adolescents, celle qu'il leur est le plus facile de parler. Dans le cas de la méthode une personne / une langue, les adolescents ne se sont pas mis d'accord sur une langue commune à cause de la présence de la chercheure, l'un parlant allemand comme d'habitude, l'autre lui répondant systématiquement en français, mais chacun a opté pour une langue et s'y est tenu. Ils n'ont ni alterné les langues ni fait usage d'emprunts, comme l'imposent les principes parentaux.

Chez d'autres adolescents, élevés selon le principe langue de la famille vs. langue de l'environnement dans une famille bilingue, la présence d'une tierce personne semble ne pas avoir eu d'influence sur la langue des frères et sœurs, puisque le français était la langue

parlée, entrecoupée toutefois de divers éléments du discours bilingue. Les disputes ont eu lieu en allemand, ainsi que les adolescentes l'avaient expliqué au cours de l'entrevue. La raison donnée pour les disputes en allemand divergeait : l'une des deux adolescentes, bilingue de cœur, a décrété qu'il y a plus de gros mots en allemand... tandis que l'autre, influencée par l'attitude plutôt négative de sa mère vis-à-vis de la culture allemande et rejetant en partie son côté allemand, déclarait : « c'est parce que les allemands ils sont brutaux ». Dans l'exemple suivant, on voit que les adolescentes, qui dans le reste du corpus se sont limitées au français, ont recours à l'allemand lorsqu'elles sont en conflit. De plus, Myriam utilise les compétences selon elle médiocres de sa sœur en allemand pour la rabaisser. Ce processus a été observé également dans une autre famille.

Myriam :	?t'arrêtes de bouder?[1]
Julia :	sûrement pas
Myriam :	ver´tragen
Julia :	non sûrement pas avec toi
Myriam :	also ?spielen wir cluedo oder nicht? + julia + ich hab dich was gefra´agt + hallo + julia + manno julia + hallo
Julia :	ja + so mein name
Myriam :	@@ xx deutsch sprechen ok
Julia :	?de quoi?
Myriam :	´wenn du deutsch sprichst dann mal richtig ?ja?
Julia :	ça se dit très bien + so mein name
Myriam :	?ja mein name?
Julia :	nein + ja so mein name habe ich gesagt
Myriam :	@du hast gesagt ja mein name@

[1] Les entrevues ainsi que les enregistrements de conversations ont été transcris à l'aide du système de transcription mis au point au sein du projet de recherche P-MOLL de Norbert Dittmar à l'Université Libre de Berlin.

Influence des principes et stratégies

Certaines différences dans le comportement linguistique des adolescents semblent clairement explicables par l'influence des principes et stratégies appliqués par leurs parents. En situation de communication familiale, le schéma imposé par les parents est reproduit à l'âge de l'adolescence. Seulement, dans le cas d'un rejet catégorique ou de tensions très importantes dans la relation parent / enfant, ce schéma semble ne pas être reproduit.

Dans cette famille où le père est allemand, on emploie la méthode langue de la famille vs. langue de l'environnement dans une famille bilingue. On ne parle que français, sauf dans quelques cas comme celui-ci, où la fille utilise le changement de langue dans le but stratégique d'amadouer son père, qui déclare vouloir participer au bal de l'école. Toute la séquence se déroule en français, sauf cette petite partie où Julia, qui ne parle quasiment jamais allemand, change soudainement de langue mais revient très vite au français :

Alexandra :	au fait dominik t'as le droit d'aller au bal + t'es invité aussi si tu veux^
Julia :	oui vous avez le droit + vous pouvez y aller si vous avez envie
Myriam :	ah oui
Dominik :	oui ben alors je viens avec vous de toute façon
Alexandra :	oui tu y vas avec eux pour les surveiller
Myriam :	ah ´nan
Dominik :	on ´danse ensemble
Julia :	non + bitte papa^ + bitte papa^
Myriam :	bleib zu hause
Julia :	du hast besseres zu tun
Myriam :	ja du arbeitest zu hause am computer
Julia :	ah non papa + attends + j'ai pas envie de me ridiculiser moi attends

L'hypothèse formulée au début des travaux n'a pas pu être confirmée, selon laquelle les adolescents reproduisaient systématiquement le schéma de l'éducation parentale. Cependant, il

est clair que ceux éduqués selon le principe une personne / une langue tendent plus à changer de langue lorsqu'ils changent d'interlocuteur que les autres. L'exemple suivant en est une illustration. La chercheure s'est adressée à Christoph en allemand, et la mère de celui-ci a montré son étonnement par rapport à ce qu'elle disait, de sorte qu'il lui a répondu, en allemand, et s'est tourné vers l'enquêteure pour lui parler, en français.

Céline :	ich will deine eltern auch interviewen
Elke :	?ja?
Christoph :	ja klar und heute abend +?tu restes jusqu'au soir?
Céline :	non
Christoph :	si parce que + ils s'en vont

Outre une application quasi-systématique de la fonction directive du changement de langue, ils évitent, même en interaction avec leurs amis, d'employer des emprunts, même si dans de telles situations, ils en utilisent plus qu'en présence de leurs parents, qui appliquent certaines stratégies dans le but de les empêcher de « mélanger » les langues : s'ils utilisent un mot allemand dans une phrase en français, ou vice-versa, en présence de leurs parents, ils doivent mettre 10 *pfennigs* dans une caisse prévue à cet effet. Le fait que les deux langues, en conséquence de l'éducation bilingue, aient été toujours utilisées de façon séparée, conduit à une forte probabilité d'interférences quand les adolescents se mettent à faire usage des divers éléments du discours bilingue. Voici quelques exemples du corpus où Marion se trouve avec son amie au café, situation dans laquelle elle utilise les deux langues dans la même conversation de manière généreuse. On note alors une influence de l'allemand sur le français plus importante que dans les situations où la séparation entre les deux langues est plus claire :

Marion :	je me suis vachement bien compris avec eux (> ich habe mich mit ihnen sehr gut verstanden)
Géraldine :	?ça fait combien de temps que tu es à l'université?
Céline :	six ans et demi
Marion :	si long^ (> so lange!)

Lorsque, au contraire, aucune méthode n'est appliquée consciemment ou que les principes des parents ne sont pas trop rigides, le danger semble moindre d'observer des interférences d'une langue sur l'autre au moment de l'adolescence, lorsque de toute façon, tous les adolescents, quelle que soit leur attitude face aux deux langues, font usage de ces divers éléments, ainsi que l'a montré l'étude des données.

Méthode d'éducation bilingue et perception

La méthode d'éducation bilingue semble avoir une influence remarquable sur la perception des adolescents, celle-ci étant plus aiguë lors d'une séparation claire des deux langues, comme dans le cas de la méthode une personne / une langue, en comparaison avec une éducation bilingue ne se fondant sur aucune stratégie ni aucun principe conscients de la part des parents. La perception des fonctions internes de la langue, telles que la pensée, le calcul, le rêve, la joie, la tristesse, etc., semble également être sous l'influence de la méthode d'éducation bilingue. Tous les informateurs ont affirmé que ces processus ont lieu dans les deux langues, sauf en cas de rejet catégorique de l'une des langues. Seule la combinaison est différente : les jeunes éduqués selon le principe du lieu disent réaliser ces divers aspects dans une langue ou dans l'autre selon le lieu où ils se trouvent. Si aucune méthode n'a été appliquée, les processus internes ont lieu, semble-t-il, en alternant d'une langue à l'autre, sans que le locuteur puisse le définir. Cependant, l'attitude de l'adolescent face à ses langues semble être un facteur décisif en ce qui concerne sa perception des fonctions internes des langues. Une attitude positive entraîne par exemple une division des tâches remplies par les différentes langues qui reflète cette attitude : tout ce qui est positif est fait dans une langue, tout ce qui est négatif dans l'autre, selon les informateurs.

La relation entre l'attitude face au bilinguisme et le comportement linguistique est apparue comme évidente : en cas d'une attitude franchement négative, l'adolescent rejette carrément l'une de ses langues. En cas d'une attitude plutôt négative, il fait usage majoritairement de l'une des langues, tout en insérant dans son discours quelques éléments de l'autre langue. En revanche, une attitude positive face au bilinguisme entraîne une utilisation systématique des divers éléments du discours bilingue. Cette attitude positive ou négative est fortement influencée par l'attitude des parents.

Attitude et comportement linguistique

Il est par contre difficile d'affirmer quelle est l'influence de l'attitude face aux deux langues et face au bilinguisme sur l'emploi de ces divers éléments. Il semble que les emprunts répandus et *tag-switchings* soient considérés comme plutôt neutres, car ils sont l'expression d'une habitude langagière qui peut apparaître dans une langue comme dans l'autre. Julia, par exemple, malgré son attitude plutôt négative, a utilisé *echt*, qui veut dire « vraiment », de manière généreuse. En revanche, l'emploi d'emprunts spontanés ne semble plus être considéré à l'âge adolescent comme négatif, contrairement à ce qui a été observé chez de jeunes enfants (Kielhöfer et Jonekeit, 1983 : 135). Dans le discours de presque tous les jeunes, on retrouve l'emploi d'emprunts spontanés, que ceux-ci se présentent également dans l'interaction familiale ou non.

Marion :	elle me dit (h) t'es trop grosse tu dois **abnehmen**
Christoph :	il était trois mètres de + trois **fussballfelder** de long et deux de large.
Julia :	elle a vu le **kellner**
Myriam :	boah ham + letztes mal als ich bei dir war da stand doch so eine **affiche** + hem hem +

L'alternance de langue est considérée par tous comme un élément positif, marqueur d'identité avec le groupe de pairs.

Fonctions des occurrences

La majorité des alternances de langues observées ont eu un rôle fonctionnel, c'est-à-dire que les adolescents les ont utilisées dans un but bien précis, même si ce but est le plus souvent inconscient. Tous en ont fait plus ou moins usage, selon la fréquence à laquelle ils les utilisent dans le cercle familial. Si depuis leur enfance, ils ont été habitués à utiliser les deux langues en alternance, il leur est plus facile à l'adolescence d'alterner les langues et d'appliquer les diverses fonctions des éléments du discours bilingue.

Tous ont exprimé leur identité bilingue au travers de leur discours. Selon le degré d'importance qu'ils accordent à cette identité, ils ont plus ou moins alterné les langues et eu recours à l'emprunt. Tous ont montré leur particularité en public en utilisant leurs deux langues.

Marion :	?hast du (xx) gesehen?
Géraldine :	?ça parle de quoi?
Marion :	c'est deux + c'est dur parce que + c'est une fille + de d'un mec qu'était riche disons + und + die wird halt entführt erstmals und die will es aber + die hasst ihren vater weil er kann das (xxx) aber + il préfère son argent à sa fille +
Géraldine :	hm
Marion :	mais elle est adoptée + (x) die zwei haben zusammen gewohnt da
Géraldine :	avec le mec qui l'a entführt

Souvent, la reproduction de discours a été prétexte à une alternance de langues. Dans l'exemple suivant, Myriam rapporte que sa petite nièce demande toujours « pourquoi ». Les paroles rapportées sont exclusivement en français, les commentaires en allemand.

Myriam : wenn sie zum beispiel fragt ehm pourquoi tu dois manger et ehm
Oriana : pourquoi je dois pas manger
Myriam : dann sagst du + parce que j'ai faim + dann sagt sie + pourquoi + et toi t'as pas faim toi + dann hört sie auf @@ +5+ hop

La fonction directive du langage a été utilisée par tous, beaucoup plus cependant par ceux éduqués selon le principe une personne / une langue, afin de signaler un changement d'interlocuteur.

La fonction référentielle de l'alternance de langues a été également utilisée par tous, surtout en cas de « trou », comme dans l'exemple suivant, où les termes *patin*, *patinoire* ne sont pas venus à l'esprit de Christoph, qui a opté pour un changement de langue, évitant ainsi l'emprunt, honni par ses parents.

Christoph : nous tous les deux + wir sind ja vom eislauf gekommen [...]

La fonction métalinguistique du changement de langue a été utilisée en particulier par Julia, qui l'a employé afin de commenter ses propres paroles. Celle-ci, qui fait un rejet partiel de l'allemand, a utilisé l'alternance de langue dans ces circonstances-là exclusivement, sauf une fois pour amadouer son père, comme mentionné plus haut.

Julia : ?on le prend? ?comment tu veux grignoter après dans le cours si t'as pas de sucre? + das war nicht witzig

La fonction poétique de l'alternance de langue est restée peu employée, ceci étant probablement dû à la difficulté qu'entraîne un tel changement de langue. Ici, le père lève son verre en disant « à la santé de Céline », en utilisant la préposition *auf*, qui signifie également « sur », et Christoph en fait une traduction littérale qui donne « sur Céline », ce qui permet à l'adolescent de jouer avec ses

deux langues, privilège partagé avec les initiés seulement, le sens du jeu de mots échappant à l'interlocuteur non-bilingue.

Stefan :	auf céline
Christoph :	sur céline
Céline :	?sur céline?
Christoph :	sur céline

Conclusion

Dans le cas d'une éducation bilingue réussie, la compétence linguistique du locuteur exerce une influence mineure sur son comportement. Les facteurs d'influence sont alors la personnalité de l'adolescent, l'histoire langagière individuelle, la méthode d'éducation bilingue appliquée par les parents, le besoin de s'identifier au groupe de pairs. L'aspect de l'identification est essentiel, car l'adolescence est un âge auquel ce besoin d'identification est particulièrement fort, ce qui conduit les jeunes à développer une langue commune qui sert de base à un sentiment de solidarité au sein d'un groupe fermé. Il semble donc inévitable qu'à l'âge de l'adolescence, les individus bilingues se mettent à faire usage des divers éléments du discours bilingue. La méthode d'éducation bilingue appliquée par les parents représente dans ce cadre-là un facteur décisif d'influence du comportement linguistique de l'adolescent. Plus l'individu est libre de changer de langue au cours de son acquisition, moins il semble avoir de difficultés à vivre sa bilingualité au moment de l'adolescence. S'il est habitué à changer de langue, il peut le faire avec plus de facilité et explorer le monde bilingue avec créativité. En revanche, il est évident que l'adolescent qui n'a pas été libre de passer d'une langue à l'autre dans l'enfance ne peut le faire avec autant d'habileté, et le nombre d'interférences augmente alors de façon remarquable lorsque les deux langues sont utilisées dans une même conversation. L'intransigeance avec laquelle les parents appliquent leurs principes détermine la liberté de l'enfant à jouer avec la

langue de façon à en appliquer les diverses fonctions, ainsi que sa faculté à les utiliser à l'adolescence.

Bibliographie

APPEL, R., et P. MUYSKEN (1987), *Language Contact and Bilingualism*, Londres, Edward Arnold.

FISHMAN, J. A. (1965), « Who Speaks What Language to Whom and When? », *Linguistique 1*, p. 67-88.

HOFFMANN, C. (1991), *An Introduction to Bilingualism*, Londres et New York, Longman.

KIELHÖFER, B., et S. JONEKEIT (1983), *Zweisprachige Kindererziehung*, Tübingen, Stauffenberg.

MAHLSTEDT, S. (1996), *Zweisprachigkeitserziehung in gemischtsprachigen Familien: eine Analyse der erfolgsbedingenden Merkmale*, Frankfurt am Main.

POPLACK, S. (1989), « Statut de langue et accommodation langagière le long d'une frontière linguistique », dans R. MOUGEON et É. BENIAK (dir.), *Le français canadien parlé hors Québec*, Québec, P.U.L., p. 127-151.

ROMAINE, S. (1989), *Bilingualism*, Oxford et New York, B. Blackwell.

Est-ce que les comparatives et les directionnelles sont des séries verbales?*

Renée Lambert
Université du Québec à Montréal

Introduction

Il existe dans plusieurs langues, notamment en fon, une langue de l'Afrique de l'Ouest, des constructions verbales dites sérielles où deux (ou plusieurs) verbes sont mis en série pour former un seul prédicat complexe. Les verbes de ce type de construction ne sont séparés ni par un marqueur de conjonction ni par un marqueur de subordination et la série verbale est interprétée le plus souvent comme un seul événement (Lefebvre, 1991). La phrase (1) en constitue un exemple.

(1) Kòkú só àsón ná Àsíbá.
Koku prendre crabe donner Asiba
« Koku a donné le crabe à Asiba ». (=(1c) dans Lefebvre, 1991)

Dans cette construction que je nomme simple parce que c'est le type de sérielle le plus étudié et à propos duquel il n'y a pas beaucoup de controverse, *Koku* est le sujet des deux verbes, *crabe* est compris comme étant à la fois l'argument interne (ou objet) du premier verbe et du deuxième verbe, et *Asiba* est l'objet du deuxième verbe seulement. On dira alors que *Koku* et *crabe* sont des arguments partagés.

*Je tiens à remercier Claire Lefebvre et Catherine Léger pour leurs judicieux commentaires. Je remercie également Lydie Agnan pour avoir partagé avec moi ses connaissances du fon. Cette recherche est possible grâce à une subvention FCAR (1997-1999) accordée à Mark Baker, Claire Lefebvre et Lisa Travis, et à une subvention spéciale de la vice-rectrice à la recherche et à la création de l'UQAM (1999) attribuée à Claire Lefebvre.

Claire Lefebvre (1991 : 73) a relevé certaines propriétés générales des verbes composant les constructions sérielles simples[1].

V1 : Verbs of the closed class (Take)
-active verb containing a causation component
-Agent
-Theme of change of location ;

V2 : Verbs of the open class
-verb of change of location lacking a causation component
-no overt Agent
-Theme of change of location.

De plus, Mark Sebba (1987) propose sept critères pour distinguer les séries verbales subordonnantes des séries verbales coordonnantes. Les voici.

A- V1 et V2 doivent être des verbes lexicaux (c'est-à-dire peuvent être des verbes seuls dans une phrase simple) ;
B- Il n'y a qu'un seul sujet syntaxique de surface ;
C- Le sujet sémantique de V2 est soit le sujet, soit l'objet de V1 ;
D- V1 et V2 apparaissent dans la même proposition (c'est-à-dire qu'il n'y a pas de marqueur de subordination ou de coordination) ;
E- Si jamais il est possible de concevoir V1 et V2 comme dénotant des actions séparées, alors V1 et V2 sont interprétés avec la même référence temporelle ;
F- La négation, qu'elle soit marquée une fois ou plus d'une fois, a portée sur l'ensemble de la série ;
G- Le temps, l'aspect, le mode et la polarité sont soit marqués une fois, soit marqués pour V2 avec la même valeur que V1.

[1] Le V1 provient d'une classe fermée de verbes, souvent *prendre*. C'est un verbe dynamique ayant un composant de causation, il nécessite un agent et son objet est un thème de changement de lieu. Le V2 provient d'une classe ouverte de verbes. C'est un verbe de changement de lieu n'ayant pas de composant de causation, son agent est implicite et son objet est un thème de changement de lieu.

Plusieurs analyses ont été proposées pour rendre compte des constructions sérielles. Mark Baker (1989, 1991), Collins (1993, 1997), Rose-Marie Déchaine (1993), entre autres, postulent que le partage de l'argument interne, soit le thème de changement de lieu selon Lefebvre, est obligatoire. Paul Agbedor (1993), Maxime da Cruz (1994), Paul Law et Tonjes Veenstra (1992), Comlan Tossa (1994), eux, tentent de montrer que le partage de l'objet n'est pas obligatoire en présentant des contre-exemples.

L'hypothèse de cette recherche est que, suivant Baker (1989, 1991), Collins (1993, 1997) et Déchaine (1993), le partage de l'argument interne est une condition essentielle à la sérialisation. Pour vérifier cette hypothèse, il importe d'étudier lesdits contre-exemples pour suggérer qu'ils n'en sont pas. Nous regarderons donc les constructions comparatives et directionnelles par rapport à la définition ci-haut et aux propriétés des séries verbales.

Analyse des contre-exemples

Il apparaîtra que les constructions comparatives et directionnelles correspondent à la définition des séries verbales énoncée précédemment. En effet, elles sont constituées de deux verbes qui ne sont séparés ni par un marqueur de conjonction, ni par un marqueur de subordination et elles forment ce qui semble être un prédicat complexe. Nous verrons dans les sections suivantes que les comparatives et les directionnelles pourraient en fait être d'autres types de constructions.

Comparatives

Michel Platt (1992) soutient que la construction comparative en fon présente les mêmes propriétés générales que les constructions sérielles simples : 1) la présence d'un seul sujet mais plus d'un verbe ; 2) l'absence d'Aspect devant V2 ; 3) l'implication d'une

classe limitée de verbes ; 4) l'ordre fixe entre V1 et V2 ; 5) l'existence des deux verbes indépendamment ; et 6) l'absence de conjonction. Ce qui est intéressant ici est non pas les similitudes entre les deux types de constructions, mais bien les distinctions. Les constructions comparatives se caractérisent par un V2 qui est un verbe de comparaison comme *dépasser*, *surpasser*, *ressembler*, *être égal à*, etc. En voici un exemple en (2).

(2) Kɔ̀kú kóló hú Àsíbá.
Koku grossir dépasser Asiba
« Koku a grossi plus qu'Asiba ».

Deux tests syntaxiques faits sur ces constructions nous laissent douter de leur statut de sérielles. Premièrement, il y a le progressif. Le progressif est formé par ɖò… wɛ̀. Ɖò, qui veut dire « être à », sélectionne une postposition qui elle sélectionne une nominalisation. La nominalisation se construit en inversant l'ordre verbe-objet. Dans les sérielles simples, ɖò apparaît devant le premier verbe et la postposition doit obligatoirement apparaître à la fin de la phrase, comme le montre la phrase (3).

(3) a. Kɔ̀kú ɖò kókló só yí àxì mɛ̀ wɛ̀.
Koku être-à poulet prendre aller marché dans POST
« Koku est en train d'apporter du poulet au marché ».
b. *Kɔ̀kú ɖò kókló só wɛ̀ yí àxì mɛ̀
Koku être-à poulet prendre POST aller marché dans

Pour les comparatives, on obtient des résultats différents. En effet, la postposition peut apparaître soit à la fin de la phrase, soit entre les deux verbes.

(4) a. Kɔ̀kú ɖò kìkló hú Àsíbá wɛ̀.
Koku être-à grossir dépasser Asiba POST
b. Kɔ̀kú ɖò kìkló wɛ̀ hú Àsíbá.
Koku être-à grossir POST dépasser Asiba
« Koku est en train de grossir plus qu'Asiba ».

Le test du progressif montre que le deuxième verbe des comparatives peut suivre la postposition. Cela laisse un doute sur son statut verbal. En effet, en mettant une construction sérielle simple au progressif, les deux verbes doivent apparaître ensemble. Ceci laisse entendre qu'il n'y a qu'un seul prédicat. Ce ne serait pas le cas des comparatives où les verbes peuvent être séparés par la postposition.

Un deuxième test est l'extraction des arguments. Dans les constructions sérielles simples, il est possible d'extraire tous les arguments.

(5) a. Kɔ̀kú wɛ̀, sɔ́ kókló yí àxì mɛ̀
Koku c'est prendre poulet aller marché dans
« C'est Koku qui a apporté du poulet au marché ».

b. Kókló wɛ̀, Kɔ̀kú sɔ́ yí àxì mɛ̀
poulet c'est Koku prendre aller marché dans
« C'est du poulet que Koku a apporté au marché ».

c. Àxì mɛ̀ wɛ̀, Kɔ̀kú sɔ́ kókló yí
marché dans c'est Koku prendre poulet aller
« C'est au marché que Koku a apporté du poulet ».

Dans le cas des comparatives, mes informateurs ont jugé agrammaticale l'extraction de l'objet du deuxième verbe, alors que les informateurs de Platt (1992) l'ont jugé grammaticale.

(6) a. Kɔ̀kú wɛ̀, kóló hú Àsíbá.
Koku c'est être-gros dépasser Asiba
« C'est Koku qui est plus gros qu'Asiba ».

b. *Àsíbá wɛ̀, Kɔ̀kú kóló hú
Asiba c'est Koku être-gros dépasser
(ok = (28) dans Platt, 1992)

L'impossibilité d'extraire l'objet du deuxième verbe tend à montrer que celui-ci forme un constituant avec son argument, ce qui n'est pas le cas pour les sérielles simples.

Premièrement, dans ce type de constructions, le V1 provient d'une classe de verbe ouverte, alors que le V2 provient d'une classe fermée. Ceci est l'inverse de ce que l'on obtient dans les constructions sérielles simples, tel que le prédit les propriétés générales de Lefebvre (1991). Deuxièmement, le premier verbe est intransitif, c'est-à-dire qu'il n'a pas d'objet, et le deuxième verbe de la série, dans ce cas le verbe de comparaison, requiert la présence d'un agent. Cette observation montre une propriété contraire à celle des constructions sérielles simples qui dit que l'agent du deuxième verbe ne peut qu'être implicite. Ce sont les raisons pour lesquelles on prétend que seul le sujet est partagé. Pourtant, cela n'est pas surprenant. En effet, en français par exemple, il existe quantité de phrases impliquant deux verbes ainsi que le partage du sujet. S'agit-il pour autant de constructions sérielles?

Directionnelles

Les directionnelles se caractérisent par un V1 de mouvement et un V2 qui indique la direction du mouvement. Baker (1989) et Vicky Carstens (1988) pour le yoruba, Agbedor (1993) pour l'éwé, Richard Campbell (1992) pour l'akan, Veenstra (1996) pour le sarramaccan, etc. Tous ont traité les constructions présentant les caractéristiques des directionnelles comme étant des cas de sérialisation. La phrase suivante en constitue un exemple.

(7) Kòkú kán-wèzùn yì kóxò.
Koku faire-course aller dehors
« Koku est allé dehors en courant. / Koku a couru pour aller dehors ».

Cet exemple présente certaines propriétés générales des séries verbales simples, à savoir qu'il n'y a qu'un seul sujet mais plus d'un verbe, qu'il n'y a que certains types de verbes impliqués, qu'il y a un ordre fixe entre les verbes de la construction, que les deux verbes existent indépendamment et qu'il n'y a ni marqueur de subordination ni marqueur de conjonction séparant les deux verbes. Mais encore une fois ici, une différence notable est qu'il n'y a pas de partage d'argument interne. Seul le sujet est partagé.

Deux tests syntaxiques ont aussi été appliqués sur ces constructions. Premièrement, le progressif. Rappelons que dans les sérielles simples, les deux verbes doivent apparaître ensemble, comme en (3). Ce n'est pas le cas pour les directionnelles.

(8) a. Kɔ̀kú ɖò wèzùn-kán yì kɔ́xò wɛ̀.
Koku être-à course-faire aller dehors POST
« Koku est en train d'aller dehors en courant ».

b. Kɔ̀kú ɖò wèzùn-kán wɛ̀ (ná) yì kɔ́xò.
Koku être-à course-faire POST FUT aller dehors
« Koku est en train de courir pour aller dehors ».

L'exemple (8b) montre qu'il est possible pour les directionnelles que le premier verbe soit au progressif et le deuxième verbe au futur. Cela est contraire aux sérielles simples où un seul marqueur T(emps)M(ode)A(aspect) apparaît devant les deux verbes.

Un deuxième test consiste à mettre la construction au négatif. La négation *ă* apparaît en fin de phrase et a portée sur les deux verbes dans le cas des simples.

(9) Kɔ̀kú sɔ́ kókló yí àxì mɛ̀ ă.
Koku prendre poulet aller marché dans NEG
« Koku n'a pas apporté du poulet au marché ».

Ce n'est pas ce qu'on retrouve pour les directionnelles. La négation a portée sur l'un ou l'autre des deux verbes, mais pas sur les deux.

(10) Kòkú kán-wèzùn yì kóxò ǎ.
Koku faire-course aller dehors NEG
« Koku n'a pas couru pour aller dehors (il a marché) ».
« Koku n'est pas allé dehors en courant (il n'est pas allé dehors du tout) ».

La possibilité d'avoir la portée de la négation sur l'un ou l'autre des deux verbes montre que les deux verbes de la construction ne forment pas un seul prédicat complexe. Les deux tests tendent à montrer que les deux verbes sont indépendants dans la phrase et ne sont pas liés comme dans les constructions sérielles simples.

Premièrement, selon les propriétés générales des constructions sérielles données par Lefebvre (1991), le premier verbe de la série proviendrait d'une classe restreinte de verbe et serait un verbe qui implique une causation. Tous les V1 des directionnelles semblent effectivement provenir d'une classe fermée de verbe (verbes de mouvement), mais les directionnelles ne rencontrent pas l'exigence de la causation. D'autre part, toujours selon Lefebvre (1991), le deuxième verbe de la série proviendrait d'une classe ouverte de verbe, mais ce n'est pas ce que l'on obtient avec les directionnelles. En effet, le deuxième verbe provient lui aussi d'une classe restreinte de verbe, en ce sens que ce doit être un verbe qui puisse orienter le mouvement. Ces caractéristiques sont différentes de ce que l'on observe pour les séries verbales simples. Comme pour les comparatives, seul le sujet est partagé dans les directionnelles. Ainsi que nous l'avons mentionné pour les comparatives, le partage du sujet n'est pas une condition suffisante pour les considérer comme étant des sérialisations.

Conclusion

Il apparaît de prime abord que les constructions comparatives et directionnelles correspondent à la définition des constructions sérielles énoncée précédemment. En effet, on observe qu'elles sont constituées de deux verbes séparés ni par un marqueur de

conjonction ni par un marqueur de subordination et on pourrait penser qu'elles forment des prédicats complexes. Cependant, à la lumière des tests syntaxiques opérés sur ces constructions, il semble plutôt que les comparatives et les directionnelles soient d'autres types de constructions. En n'étant pas des séries verbales, elles ne constituent pas des contre-exemples à mon hypothèse, qui est que le partage de l'argument interne par les verbes de la série est une condition à la sérialisation. Par contre, pour les constructions sérielles simples, il y a un partage de l'objet.

Le but de cet article a été non pas de proposer une classification solide de ces deux types de constructions dites sérielles en fon, mais plutôt de suggérer qu'en général, on classe peut-être un peu trop vite comme étant des séries verbales les constructions qui présentent certaines propriétés des véritables constructions sérielles, le plus souvent le partage d'argument, externe et / ou interne. Il reste encore beaucoup de travail à faire dans l'étude des constructions phrastiques de tous ordres en fon. En effet, peu d'auteurs se sont penchés sur la véritable nature des différentes phrases en fon. Le présent article constitue un ensemble de pistes à suivre ou à falsifier pour arriver à une classification plus juste des constructions sérielles, et ce, pour toutes les langues où le phénomène a été observé.

Bibliographie

AGBEDOR, P. (1993), « Verb serialization in Ewe », *Studies in the Linguistic Sciences 23*, p. 21-42.

BAKER, M. (1991), « On the relation of serialization to verb extensions », dans C. LEFEBVRE (éd.), *Serial Verbs : Grammatical, Comparative and Cognitive Approaches*, Amsterdam, John Benjamins, p. 79-102.

BAKER, M. (1989), « Object sharing and projection in serial verb constructions », *Linguistic Inquiry, 20, 4*, Cambridge (MA), MIT Press, p. 513-554.

CAMPBELL, R. (1992), « Serial verbs and unaccu: ativity », *MIT Working Papers in Linguistics, 17*, p.41-51.

CARSTENS, V. (1988), *Serial Verbs in Yoruba*, paper presented at the Second Workshop in Niger-Congo Syntax and Semantics, Cambridge (MA), MIT.

DA CRUZ, M. (1994), « Les constructions sérielles en fɔngbe : approches sémantique et syntaxique », *Travaux de recherche sur le créole haïtien 20-21*, Montréal, UQAM.

DÉCHAINE, Rose-Marie (1993), « Serial verb constructions », dans J. JACOBS, A. VON STECHOW, W. STERNEFELD et T. VENNEMANN (éd.), *Syntax : An International Handbook of Contemporary Research*, Berlin, Walter de Gruyter, p. 799-825.

LAW, P., et T. VEENSTRA (1992), « On the structure of serial verb constructions », *Linguistic Analysis 22, 3-4*, New York, American Elsevier Pub., p.185-217.

LEFEBVRE, C. (1991), « Take serial verb constructions in Fon », dans C. Lefebvre (éd.), *Serial Verbs : Grammatical, Comparative and Cognitive Approaches*, Amsterdam et Philadelphie, John Benjamins, p. 37-78.

PLATT, M. M. (1992), « Les constructions comparatives en fongbe, un cas de sérialisation? », *Travaux de recherche sur le créole haïtien 7*, Montréal, UQAM, p. 91-97.

SEBBA, M. (1987), *The Syntax of Serial Verbs : An Investigation into Serialisation in Sranan and Other Languages*, Amsterdam et Philadelphie, John Benjamins.

TOSSA, C. Z. (1994), « Adjonctions et séries verbales dans les langues GBE ». Thèse de doctorat, Ottawa, Université d'Ottawa.

VEENSTRA, T. (1996), *Serial Verbs in Saramaccan. Predication and Creole Genesis*, Dordrecht, ICG.

Langue et identité: perspectives des Inuits*

Shelley Tulloch
Université Laval

Introduction

La question de l'importance d'une langue ancestrale pour l'identité ethnique est pertinente chez les Inuits, un peuple qui fait face actuellement au déclin dans l'usage de leur langue ancestrale. Les tentatives de démontrer un lien entre la langue et l'identité inuites de l'extérieur donnent des résultats limités car il est difficile pour les étrangers de cerner adéquatement à la fois l'essentiel de l'identité inuite et les subtilités de l'inuktitut, une variété de la langue inuite. Quand les Inuits parlent de l'importance de l'inuktitut pour leur identité, ils révèlent d'abord de la perspicacité concernant ces deux éléments fondamentaux à leur être. Ils dévoilent, en outre, les motivations qui peuvent sous-tendre certains comportements et leur probabilité d'entreprendre d'autres actions. Cet article présente la voix des Inuits, révélant leurs perceptions de l'importance de la langue inuite pour l'identité inuite. Leurs commentaires tendent à démontrer que l'inuktitut constitue, en effet, une partie importante de l'identité inuite, et que cette importance dépend largement du rôle que joue l'inuktitut dans la société inuite.

Méthodologie

Cette discussion sur la langue et l'identité inuites se base sur les réponses obtenues au cours d'entrevues semi-dirigées effectuées auprès d'un échantillon de la population inuite à Iqaluit, capitale du Nunavut. Ces entrevues, qui s'insèrent dans la méthodologie d'un

* Nous tenons à remercier le ministère des Affaires indiennes et du Nord et le gouvernement de l'Ontario qui ont subventionné le travail de terrain de ce projet.

plus vaste projet visant l'évaluation des possibilités d'un aménagement linguistique au Nunavut, invitent les participants à s'exprimer sur une variété de thèmes, dont les choix quotidiens d'usage linguistique, les motivations qui sous-tendent le choix d'employer une langue plutôt qu'une autre et l'importance de l'inuktitut. Les réponses de trois informateurs représentatifs ont été sélectionnées pour cette analyse.

Ces informateurs sont inuits, de parents inuits. Ils parlent tous couramment l'inuktitut, leur langue maternelle, et jouissent d'une bonne compétence en anglais aussi. Ils ont tous habité à Iqaluit pendant plusieurs années et chacun a aussi habité ailleurs, dans le sud du Canada ou dans d'autres communautés inuites. Les informateurs ont tous au moins entamé les études post-secondaires. Deux des informateurs sont des femmes, l'autre est un homme. Ils représentent les trois tranches d'âge ciblées dans cette étude : 20-29, 30-39 et 40-49. L'un des informateurs est étudiant au Collège Arctique, un autre est cadre supérieur dans une organisation inuite et le dernier est ingénieur stationnaire. Pendant les entrevues dirigées en anglais, chaque participant fait allusion à l'identité en parlant de l'inuktitut et du choix d'usage linguistique.

Discussion

Deux thèmes principaux concernant l'importance de l'inuktitut pour l'identité inuite ressortent d'une analyse compréhensive des commentaires directs et indirects des informateurs. En premier lieu, les informateurs semblent être conscients du fait que l'inuktitut est un facteur déterminant de leur identité individuelle et collective en tant qu'Inuits, à la fois parce qu'il l'est aujourd'hui (exprimé comme présomption) et parce qu'il l'a été par le passé. En deuxième lieu, et cet aspect de l'analyse est le plus porteur quand on soumet les entrevues à l'analyse, les données révèlent que l'inuktitut est important pour l'identité inuite car il est un outil, une ressource, pour créer, développer et maintenir l'identité inuite.

Toutefois, avant de procéder à l'analyse compréhensive, il est pertinent de considérer ce que les informateurs disent quand ils parlent en toute conscience de l'importance de l'inuktitut pour l'identité.

Pendant toutes les entrevues, les informateurs font des commentaires qui révèlent en quoi leur langue ancestrale est importante pour leur identité ethnique. Questionné au sujet de l'importance de l'inuktitut pour l'identité inuite, l'informateur A1 parle de l'importance de l'inuktitut pour la communication, donc pour l'entretien du réseau social :

> *[What if you didn't speak Inuktitut at all, what would that mean to you as being Inuit?]*
>
> *I think that's becoming more and more apparent. There are a lot of beneficiaries, Inuit beneficiaries* [of the Nunavut Land Claims Agreement], *who do not speak Inuktitut. It's new, you know. They are Inuit, but they don't speak it. I think, for the adults, it's hard for them to communicate. Then the communication becomes a problem. There's not enough, they're not communicating with their elders or their extended family members who are older, like older aunts or uncles or grandparents. ...There's no interaction. The same goes with the workplace when they don't speak Inuktitut, even if they understand it. Then they're missing out on something that connects you to the social structure, and as well as culture* (A1, août 1999).

L'informatrice A2, quant à elle, commence sa réponse en disant que parler l'inuktitut est un élément essentiel de l'identité inuite. Toutefois, en réfléchissant sur la question, elle adoucit sa position :

> *[If you didn't speak Inuktitut, would you still be Inuk?]*
> *I don't think so. Maybe. But language is a big part of being Inuk in our tradition. ...There was a person, and she said she was an Inuk, but she didn't speak the language. And I said, but to be an Inuk, you need to speak Inuktitut, because that's a part of the tradition. I don't know now if I sounded right.*

> [...] *I don't know. If their parents were Inuit, then I guess I would consider them, in some ways, Inuit. I guess, biologically they are Inuks.*
>
> [...] *The more I think about it, Inuks are not so much language, it's the culture. The language doesn't necessarily make up the culture, so it's part of it. But I based it on just the language, but there's other things now, like hunting, having [raw meat] to base it on.*
>
> [...] *[Language] is a big thing, but it's not everything. I think one of the most important parts of being Inuit is, it's always been, the surviving in the cold, and they had to help each other out, and that's stronger. So it's one of the stronger bases than Inuktitut* (A2, août 1999).

L'informateur A3 exprime également un fort lien entre l'inuktitut et l'identité inuite, en constatant que parler l'inuktitut est essentiel pour l'identité inuite et, qui plus est, que l'importance de parler l'inuktitut se base sur les besoins communicatifs, en particulier pour la communication avec les aînés :

> *[What if people can't speak the language, what will that mean to them?]*
>
> *It's like you lose your identity. You just lose who you are, I guess. [What's the difference between an Inuk who speaks the language and an Inuk who doesn't?]*
>
> *I don't know. He can't speak it.*
>
> *[What about if you only spoke English, what would that mean to you?]*
>
> *It would mean... talking to the elders. That's about it* (A3, octobre 1999).

Ces citations révèlent la relation que les informateurs perçoivent consciemment entre la langue et l'identité inuites. Une analyse compréhensive de tous leurs commentaires renforce les positions exprimées dans ces citations et suggère aussi d'autres raisons pour

lesquelles la langue inuite est importante pour l'identité inuite à Iqaluit.

La première notion qui ressort de l'analyse compréhensive est que l'inuktitut est en soi important pour l'identité inuite. Les informateurs expriment ce lien conscient et significatif de deux manières. D'abord, ils expriment ce lien en tant que présomption, sans y apporter de preuves : « Inuktitut… it's part of our identity in some ways. It's part of who we are » (A2, août 1999), « [If you lose your language] It's like you lose your identity. You just lose who you are » (A3, Août 1999), « In any presentation that I have to do, I'm much more comfortable doing it in Inuktitut, because that's what makes me who I am » (A1, août 1999). Tous les informateurs affirment que l'inuktitut est important pour l'identité inuite. Pourtant, ils donnent relativement peu de preuves afin d'appuyer cette affirmation.

La seule preuve d'un lien absolu et inhérent entre l'inuktitut et l'identité inuite qui ressort des données est reliée au fait que l'inuktitut est une partie importante du passé et des traditions inuites : « To be an Inuk, you need to speak Inuktitut, because that's part of the tradition » (A2, août 1999). L'informateur A3 explique la même idée d'une autre manière, en affirmant que savoir parler l'inuktitut est primordial, tandis que savoir l'écrire l'est moins : « As long as they can speak it. The written language is not even ours. It's not our language. It's not our written language. We never had one. As long as they can speak it » (A3, octobre 1999). L'inuktitut est important pour l'identité inuite car il constitue un élément du passé et de la tradition : il incarne et rappelle ce que les Inuits ont été.

La deuxième idée qui ressort de l'analyse compréhensive est que le rôle que joue l'inuktitut dans la société, plutôt que des caractéristiques intrinsèques à cette langue, lui attribue son importance pour l'identité. Les Inuits se servent de l'inuktitut pour marquer leur identité en tant qu'Inuits ainsi que celle de leurs

interlocuteurs. L'inuktitut sert à inclure ou à exclure d'autres membres de la communauté. Aussi, cette langue aide à forger et à maintenir le réseau social. Certes, la langue sert également à créer, développer et maintenir l'identité d'autres manières aussi, mais les trois fonctions mentionnées plus haut sont les plus significatives d'après cette analyse.

Quand on parle en inuktitut à Iqaluit, on marque l'identité. Un Inuk bilingue peut choisir de parler en inuktitut afin de démontrer son identité inuite. Pour certains, il est important de parler en inuktitut afin de se caractériser comme Inuk : « He wears all the seal clothing and all that to make himself look Inuk. Well, you want to look like one but you won't try to speak it. They're the ones that make me mad » (A3, octobre 1999). Un Inuk bilingue peut également choisir de parler en inuktitut afin de démontrer qu'il reconnaît l'identité inuite de son interlocuteur. En effet, le facteur qui semble exercer la plus grande influence sur le choix de la langue est l'identité ethnique de l'interlocuteur (rappelons que l'échantillon ne comprend ni les moins de 20 ans ni les plus de 50 ans). Les Inuits parlent avec d'autres Inuits en inuktitut : « I know the town hall, the reception there, it's an Inuit person. I'll speak Inuktitut with her » (A2, août 1999). Qui plus est, parler en inuktitut avec un autre Inuk (ou une autre personne qui parle l'inuktitut) met en valeur la similitude entre ces deux personnes. L'informateur A3 a exprimé à quel point il est important pour lui qu'un Inuk lui parle en inuktitut, même si la personne n'est pas à l'aise dans cette langue : « We just told him, "You're Inuk. You'd better speak Inuktitut to us" » (A3, octobre 1999). La langue inuite est un marqueur important de l'identité à Iqaluit, à la fois pour marquer sa propre identité et celle de son interlocuteur, et pour marquer la similitude entre deux personnes.

Si parler en inuktitut est un moyen de montrer la similitude entre les locuteurs, il sert aussi à renforcer les liens d'appartenance, ou, au contraire, pour exclure de son groupe ceux qui ne parlent pas l'inuktitut. Comme l'informatrice A1 l'affirme, parler l'inuktitut a

des retombées sur sa participation dans la communauté inuite et sur le sentiment d'appartenir complètement à cette communauté :

> *But that's part of the social and culture again. I think that [those who don't speak Inuktitut] get discouraged and feel like they're not really part of that, unfortunately. So when they say, « So and so, she's Inuk, but she doesn't know how to speak Inuktitut, so she's not really Inuk », that's what they're meaning. ...Social acceptance, we all want to belong to something*
> (A1, août 1999).

De la même manière, on utilise l'inuktitut de façon consciente pour exclure d'autres personnes, en empêchant ces personnes de comprendre ce qui se passe, donc en les empêchant de communiquer et de participer. L'informateur A3 le dit clairement : « We speak in Inuktitut when we don't want other people to understand » (A3, octobre 1999). Bref, à Iqaluit, une communauté plurilingue, l'acte de parler l'inuktitut démarque ceux qui appartiennent à la communauté inuite et ceux qui n'y appartiennent pas, renforce ou nie les sentiments d'appartenance et favorise ou empêche la participation totale dans cette communauté.

Étant donné que parler l'inuktitut favorise ou empêche la communication, l'inuktitut devient donc un outil pour forger et maintenir les amitiés et les autres relations entre les résidents d'Iqaluit qui parlent l'inuktitut. Chacun des informateurs mentionne que parler en inuktitut est essentiel pour communiquer avec les aînés : « If it's an elder, for sure [I speak in Inuktitut] » (A2, août 1999). En outre, pour ceux qui ne parlent pas l'inuktitut, il est plus difficile de communiquer avec certains membres de la communauté. De là, le réseau social devient moins fort : « [Not speaking Inuktitut] would mean… [not] talking to the elders » (A3, octobre 1999), « They are Inuit but they don't speak it. I think, for the adults, it's hard for them to communicate. …They're not communicating with their elders or their extended family members who are older… » (A1, août 1999). Les relations sont un élément important pour établir et maintenir l'identité. Pourtant, la

diminution de l'usage de l'inuktitut accompagnée de l'augmentation de l'emploi de l'anglais peut menacer ces relations. Cette menace à l'identité inuite surgit peut-être à cause de la rapidité avec laquelle l'inuktitut cède sa place à l'anglais. À Iqaluit, on peut avoir les grands-parents unilingues inuktitut et les petits-enfants unilingues anglais. De cette manière, la rapidité de la perte de l'inuktitut peut occasionner des difficultés pour forger et maintenir les relations entre les personnes unilingues des deux générations.

Retombées de la perte de la langue

Les commentaires tels que : « Si on perd notre langue, on perd notre identité » et « Notre langue, c'est notre identité » sont fréquents dans les communautés inuites. Même si de tels commentaires reflètent un aspect du lien entre la langue et l'identité du point du vue des Inuits, ils transmettent une image incomplète. En fait, l'analyse précédente révèle que, lorsqu'on invite les Inuits à s'exprimer sur l'importance de l'inuktitut pour l'identité inuite, leurs commentaires directs et indirects tendent à exprimer une relation pragmatique et qualifiée entre la langue et l'identité ; cette relation va à l'encontre d'une relation inhérente et absolue suggérée par l'énoncé « notre langue, c'est notre identité ».

En effet, en disant « si on perd notre langue, on perd notre identité », on suggère que le seul élément constitutif de l'identité est la langue, comme si le schéma suivant était complet :

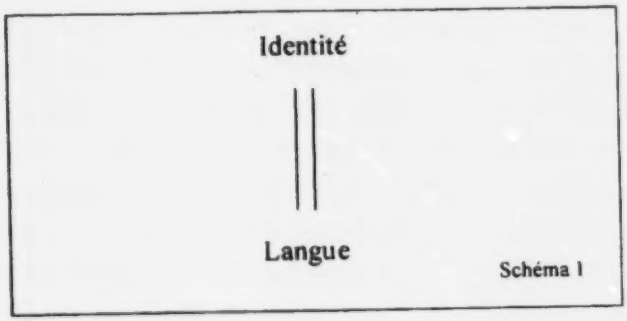


Schéma 1

Pourtant, les informateurs mentionnent plusieurs autres éléments constitutifs de l'identité inuite : la race, la survie sur la terre, la religion, le rôle des aînés comme enseignants, le réseau social, le partage, la culture, la tradition, le sentiment d'appartenance, la nourriture, pour n'en nommer que quelques-uns. Le schéma plus haut tient seulement si on a déjà perdu tous les autres aspects de l'identité inuite et que seule la langue reste. Le schéma ci-dessous reflète mieux la réalité à Iqaluit, en incluant quelques-uns des nombreux éléments constitutifs de l'identité inuite actuelle :

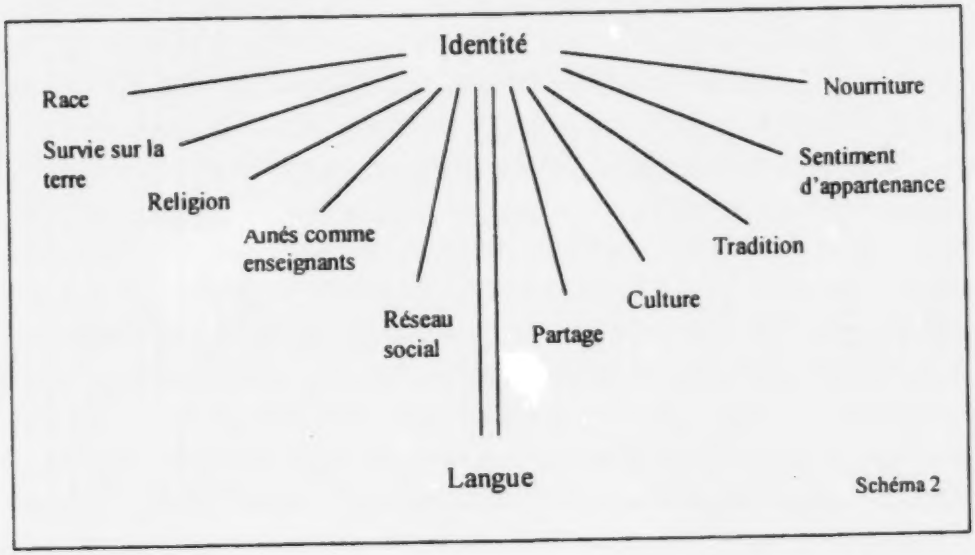


Schéma 2

Évidemment, la langue n'est pas le seul élément constitutif de l'identité inuite. Même si un seul Inuk, ou même tous les Inuits perdaient leur langue, les autres éléments culturels demeureraient.

Toutefois, les données révèlent que l'inuktitut joue un rôle important pour maintenir et soutenir d'autres éléments de l'identité. Il apparaît donc que l'importance de l'inuktitut pour l'identité inuite se trouve moins dans la langue en tant que telle et plus dans ses fonctions pratiques dans la société. La perte de l'inuktitut n'entraînerait pas nécessairement la perte de l'identité, mais représenterait quand même la perte d'un aspect important de l'identité inuite et la perte d'un pilier des autres éléments constitutifs.

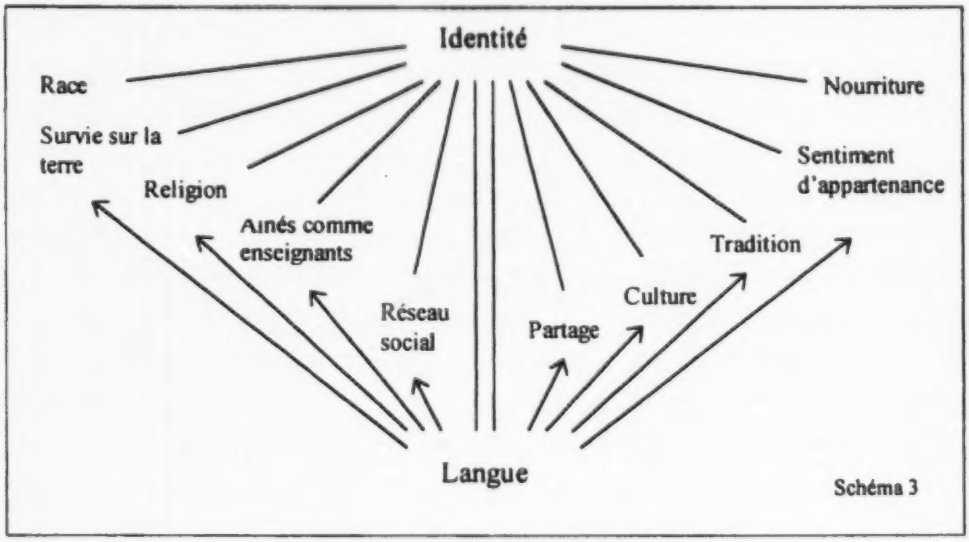


Schéma 3

Ainsi, la conséquence de la perte de l'inuktitut se trouve dans le rôle que joue l'inuktitut pour soutenir d'autres éléments de l'identité. À titre d'exemple, bien que la langue et la tradition soient deux éléments distincts, l'inuktitut appartient à la tradition et les aînés utilisent l'inuktitut pour transmettre les traditions inuites aux plus jeunes. La perte de l'inuktitut représenterait alors la perte d'un composant de la tradition ainsi que d'un outil pour le transmettre. La même chose est vraie pour la culture. De la même manière, la

perte de l'inuktitut ne détruirait pas en elle le réseau social et n'éliminerait pas nécessairement le sentiment de similitude entre les Inuits et le sentiment de différence avec les autres groupes ethniques, mais elle enlèverait un outil dont on se sert actuellement pour créer et maintenir ces aspects de l'identité inuite. Donc, l'importance de l'inuktitut pour l'identité inuite se trouve, en grande partie, dans ses fonctions pratiques qui soutiennent les autres aspects de l'identité.

Application

La perspective inuite de l'importance de l'inuktitut est signifiante et a des conséquences évidentes à Iqaluit. La croyance qu'un *vrai* Inuk doit parler l'inuktitut a des retombées positives et négatives dans cette communauté. Une telle croyance favorise, par exemple, la préservation et la promotion de la langue. Quand on valorise l'inuktitut, en affirmant l'importance de le parler afin de garder son identité inuite, on augmente chez les individus l'engagement pour préserver, protéger et promouvoir l'inuktitut. En effet, la promotion de l'inuktitut au Nunavut pourrait commencer avec une telle affirmation de l'importance de l'inuktitut pour les Inuits.

En même temps, il est important de prendre en compte la discussion de l'importance de l'inuktitut pour l'identité inuite, car la croyance d'un lien absolu entre la langue et l'identité inuite peut avoir des retombées négatives dans certaines situations. L'affirmation qu'il faut parler l'inuktitut pour être un vrai Inuk a donné lieu au cas de dénigrement et d'aliénation des personnes d'origine inuite qui ne parlent pas l'inuktitut : « Some people say you're not proud of who you are [when you don't speak Inuktitut] ... kind of ruins things when they say you're not proud » (A3, octobre 1999), « I used to get teased a lot when I got back and I only spoke English. ...I felt pretty bad » (A3, octobre 1999). En parlant de cas où les compétences restreintes en inuktitut empêchent certains Inuits de participer entièrement, l'informateur

A1 explique que de telles situations peuvent entraîner de la friction et du mécontentement : « We all want to belong to something. And I think that's where sometimes there's a bit of friction. Not friction, but discontent » (A1, août 1999). Dans d'autres cas, on met volontairement les individus qui ne parlent pas l'inuktitut à l'écart, pour nier leur appartenance au groupe et pour remettre en question leur identité inuite. Évidemment, de tels comportements n'aident guère une situation où l'on retrouve déjà souvent des crises d'identité et un manque d'estime de soi. Ces commentaires révèlent quelques conséquences négatives de l'adhérence stricte à la croyance que la langue est essentielle pour l'identité.

Conclusion

La langue est sans aucun doute importante pour l'identité. Les données montrent que les informateurs considèrent l'inuktitut comme un élément important de l'identité inuite. L'analyse des commentaires directs et indirects des informateurs inuits tendent à démontrer que l'importance de l'inuktitut pour l'identité inuite réside dans le rôle pratique que cette langue joue dans la communauté. L'inuktitut, marqueur d'identité, sert donc à inclure ou à exclure d'autres membres de la communauté. Aussi, cette langue aide à forger et à maintenir le réseau social. La réflexion sur les répercussions de la perte de la langue sur l'identité suggère que la perte de la langue n'entraînerait pas nécessairement la perte de l'identité ethnique d'un individu ou d'un peuple. Toutefois, elle représenterait la perte d'un élément important de l'identité, celui-ci étant un support à plusieurs autres éléments constitutifs de l'identité inuite. Les données suggèrent également des retombées positives et négatives quant à la croyance qui veut que la langue et l'identité jouissent d'un lien absolu, et mettent en valeur l'importance de considérer et de comprendre les perspectives des Inuits à l'égard de leur langue et de leur identité.

Bibliographie

DORAIS, L.-J. (1996a), « Inuugatta inuulerpugut: Kalallit and Canadian Inuit identities », dans B. JACOBSEN (éd.), *Cultural and social research in Greenland 95/96: Essays in honour of Robert Petersen*, Nuuk, Ilisimatusarfik / Atuakkiorfic A / S, p. 28-33.

DORAIS, L.-J. (1996b), *La parole inuit: Langue, culture et société dans l'Arctique nord-américain*, Leuven (Belgique) et Paris, Peeters.

ERIKSSON, B. (1998), « Iqaluit: Diglossia and language identity among the Inuit ». Mémoire de maîtrise non publiée, Växjö, Université de Växjö.

KWACHKA, P. (1991), « Discourse structures, cultural stability, and language shift », *International Journal of the Sociology of Language, 93*, New York, Mouton, p. 67-74.

TULLOCH, S., et C. OUELLON (2000), « Une exploration du lien entre la langue, la culture et l'identité », *Actes des XIII^es^ Journées de linguistique*, Québec, CIRAL, p. 217-226.

L'Association des étudiants diplômés inscrits en langues et linguistique **(AÉDILL)** *de l'Université Laval organise chaque année, depuis maintenant quatorze ans, un colloque permettant aux étudiants-chercheurs de deuxième et troisième cycles de présenter l'état d'avancement de leurs travaux. Il s'agit d'un des colloques étudiants les plus importants dans le domaine de la linguistique. La diversité des communications et la provenance internationale des participants en font un instrument de diffusion de premier plan de la recherche actuelle en linguistique.*

FL026847

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

REPRODUCTION BASIS

☒ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").